



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics

Citation for published version:

Leimkuhler, B, Matthews, C & Stoltz, G 2016, 'The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics', *IMA Journal of Numerical Analysis*, vol. 36, no. 1, pp. 13-79.
<https://doi.org/10.1093/imanum/dru056>

Digital Object Identifier (DOI):

[10.1093/imanum/dru056](https://doi.org/10.1093/imanum/dru056)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

IMA Journal of Numerical Analysis

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics

BENEDICT LEIMKUHLER[†], CHARLES MATTHEWS[‡]

*University of Edinburgh, School of Mathematics,
James Clerk Maxwell Building, Edinburgh, EH9 3JZ, UK*

AND

GABRIEL STOLTZ[§]

*Université Paris-Est, CERMICS (ENPC),
INRIA, F-77455 Marne-la-Vallée, FRANCE*

[Received on 9 October 2013]

We consider numerical methods for thermodynamic sampling, *i.e.* computing sequences of points which are distributed according to the Gibbs-Boltzmann distribution, using Langevin dynamics and overdamped Langevin dynamics (Brownian dynamics). A wide variety of numerical methods for Langevin dynamics may be constructed based on *splitting* the stochastic differential equations into various component parts, each of which may be propagated exactly. Each such method may be viewed as generating samples according to an associated invariant measure that differs from the exact canonical invariant measure by a stepsize-dependent perturbation. We provide error estimates à la Talay-Tubaro on the invariant distribution for small stepsize, and compare the sampling bias obtained for various choices of splitting method. We further investigate the overdamped limit and apply the methods in the context of driven systems where the goal is sampling with respect to a non-equilibrium steady state. Our analyses are illustrated by numerical experiments.

Keywords: Langevin dynamics; Stochastic differential equations; Numerical discretization.

1. Introduction

A fundamental purpose of molecular simulation is the computation of macroscopic quantities, typically through averages of functions of the variables of the system with respect to a given probability measure μ . We consider systems described by a separable Hamiltonian

$$H(q, p) = V(q) + \frac{1}{2} p^T M^{-1} p,$$

where q and p are vectors of positions and momenta, respectively, V is a potential energy function and M is a positive definite mass matrix. In the most common setting, the probability measure corresponds to the canonical ensemble. Its distribution is defined by the Boltzmann-Gibbs density, which models the configurations of a conservative system in contact with a heat bath at fixed temperature T :

$$\mu(dq dp) = Z^{-1} e^{-\beta H(q, p)} dq dp, \quad (1.1)$$

where $\beta^{-1} = k_B T$ with k_B Boltzmann's constant and Z is a normalization constant. In nonequilibrium models, where a given system is subject to nonconservative driving and dissipative perturbations, the averages may be taken with respect to a stationary distribution which has no simple functional form. Numerically, the high-dimensional integrals are approximated as ergodic averages along discrete stochastic paths (Markov chains) constructed through numerical solution of certain stochastic differential equations (SDEs).

There are two principle sources of approximation error in such computations: (i) systematic bias (or *perfect sampling bias*) related to the use of a discretization method for the SDEs (and usually proportional

[†]Email: b.leimkuhler@ed.ac.uk

[‡]Email: charles.matthews@ed.ac.uk

[§]Corresponding Author. Email: stoltz@cermics.enpc.fr

to a power of the integration stepsize Δt), and (ii) statistical errors, due to the finite lengths of the sampling paths involved and the underlying variance of the random variables. In this article we are concerned with the systematic bias, specifically the systematic bias in long-term simulation, i.e. with respect to the invariant (or nonequilibrium steady-state) distribution.

One of the most popular choices of SDE system for sampling purposes is Langevin dynamics, which is a reliable and flexible choice for both equilibrium and nonequilibrium cases. For equilibrium thermodynamics, the Langevin equations take the form:

$$\begin{cases} dq_t = M^{-1} p_t dt, \\ dp_t = -\nabla V(q_t) dt - \gamma M^{-1} p_t dt + \sqrt{\frac{2\gamma}{\beta}} dW_t, \end{cases} \quad (1.2)$$

where dW_t represents the infinitesimal increment of a standard Wiener process. The real number $\gamma > 0$ is a free parameter which may be adjusted to enhance sampling efficiency. Under suitable conditions, the dynamics (1.2) is ergodic for the Boltzmann-Gibbs distribution (see for instance Talay (2002); Mattingly *et al.* (2002); Cancès *et al.* (2007)).

The aim of this work is to provide a numerical analysis of the perfect sampling bias in Langevin dynamics arising from numerical schemes obtained by a splitting strategy, building on studies such as Talay (2002); Bou-Rabee & Owhadi (2010) and clarifying the sampling properties of recently proposed schemes Skeel & Izaguirre (2002); Melchionna (2007); Bussi & Parrinello (2007); Thalmann & Farago (2007); Leimkuhler & Matthews (2013a). Of particular interest is the behavior of methods in the overdamped limit $\gamma \rightarrow +\infty$ and variations of Langevin dynamics incorporating nonequilibrium forcings such as the addition of non-gradient forces (in which case the invariant measure is unknown). Splitting schemes based on a symplectic integration of the Hamiltonian part of the dynamics can be combined with an exact treatment of the fluctuation-dissipation part; these methods are more convenient to implement in molecular simulation codes than the implicit schemes proposed in Talay (2002); Mattingly *et al.* (2002) and are also efficient in practice Leimkuhler & Matthews (2013b). Some essential elements on the numerical analysis on the accuracy of such splitting schemes have been provided in Bou-Rabee & Owhadi (2010). We note that alternative sampling strategies are available: the bias in the invariant measure sampled by discretization of Langevin dynamics could in principle be eliminated by employing a Metropolis-Hastings procedure Metropolis *et al.* (1953); Hastings (1970) (see in particular the discussion in (Lelièvre *et al.*, 2010, Section 2.2)), but such a correction corrupts the dynamical properties of the system Bou-Rabee & Vanden-Eijnden (2009), may be costly or complicated to implement, and anyway cannot be used in situations when the invariant measure is not known; for this reason Langevin dynamics remains one of the most popular tools for molecular sampling.

We focus in this article on the case where the position space is compact (e.g. a torus, $q \in \mathcal{M} = (L\mathbb{T})^{dN}$) since this is most relevant from the point-of-view of applications in condensed matter physics and biology, where periodic boundary conditions are typically used. This assumption simplifies the treatment of the Fokker-Planck operator associated to Langevin dynamics, and, with additional smoothness assumptions on the potential energy function, ensures regularity properties, discrete spectrum and spectral gap. In particular (1.1) is the unique invariant probability measure of the Langevin process. We denote by $\mathcal{E} = \mathcal{M} \times \mathbb{R}^{dN}$ the state space.

Let us emphasize that we expect our results to hold for unbounded spaces. The proofs may however require non-trivial modifications, using in particular the tools and the results from Mattingly *et al.* (2002); Talay (2002); Bou-Rabee & Owhadi (2010).

In practice, Langevin dynamics is discretized, and averages that are computed along a single trajectory converge to averages with respect to a measure $\mu_{\gamma, \Delta t}$, which is an approximation to μ in the sense that there exists a function $f_{\alpha, \gamma}$ for which

$$\int_{\mathcal{E}} \psi(q, p) \mu_{\gamma, \Delta t}(dq dp) = \int_{\mathcal{E}} \psi(q, p) \mu(dq dp) + \Delta t^\alpha \int_{\mathcal{E}} \psi(q, p) f_{\alpha, \gamma}(q, p) \mu(dq dp) + O(\Delta t^{\alpha+1}), \quad (1.3)$$

see Section 2.4 for precise statements. Of course, the momenta are usually trivial to sample since they are distributed according to a Gaussian measure. The primary issue is therefore is to sample positions

according to the marginal of the canonical measure:

$$\bar{\mu}(\mathrm{d}q) = \tilde{Z}^{-1} e^{-\beta V(q)} \mathrm{d}q. \quad (1.4)$$

Denoting by $\bar{\mu}_{\gamma, \Delta t}$ the marginal of the invariant measure for the numerical scheme, and by

$$(\pi\varphi)(q) = \int_{\mathbb{R}^{dN}} \varphi(q, p) \kappa(\mathrm{d}p), \quad \kappa(\mathrm{d}p) = \left(\frac{2\pi}{\beta}\right)^{-dN/2} \sqrt{\det(M)} \exp\left(-\frac{\beta p^T M^{-1} p}{2}\right) \mathrm{d}p, \quad (1.5)$$

the partial average of a function φ with respect to the momentum variable, the error estimate (1.3) reads, for observables which depend only on the position variable,

$$\int_{\mathcal{E}} \psi(q) \bar{\mu}_{\gamma, \Delta t}(\mathrm{d}q) = \int_{\mathcal{E}} \psi(q) \bar{\mu}(\mathrm{d}q) + \Delta t^\alpha \int_{\mathcal{E}} \psi(q) (\pi f_{\alpha, \gamma})(q) \bar{\mu}(\mathrm{d}q) + O(\Delta t^{\alpha+1}).$$

We focus in this article on first- and second-order splitting schemes, relying on Lie-Trotter and Strang decompositions. This restriction is motivated both by pedagogical purposes and by the dominant role in applications played by second-order splitting schemes. Let us however emphasize that most of our results could be extended to higher-order decompositions.

Results corresponding to discretizations of the equilibrium Langevin dynamics and computation of static average properties are gathered in Section 2, while nonequilibrium systems and the computation of transport properties are discussed in Section 3 (illustrating the approach by the computation of the mobility or autodiffusion coefficient). The proofs of all our results can be read in Section 4.

Let us now highlight some of our contributions.

- In the equilibrium setting, we rigorously ground the results presented in Leimkuhler & Matthews (2013a) giving the leading order correction to the invariant measure with respect to Δt for general splitting schemes, via a Talay-Tubaro expansion Talay & Tubaro (1990) (see Section 2.4). We carefully study all possible splitting schemes, taking advantage of what we call the “TU lemma” (Lemma 2.4) to relate invariant measures of splitting schemes where the elementary dynamics are integrated in different orders. From a technical viewpoint, our proofs are a variation from the standard way of establishing similar results since we use the specific structure of splitting schemes to conveniently write evolution operators as compositions of elementary semigroups (working at the level of generators, as in Debussche & Faou (2012); see also Mattingly *et al.* (2010) for a related approach based on solution of appropriate Poisson equations).
- We show in Section 2.5 how the leading order correction to equilibrium averages can be estimated on the fly by approximating a time-integrated correlation function. This can be seen as a practical way of numerically solving a Poisson equation (a standard way of proceeding when studying linear response of nonequilibrium systems) and is an alternative to Romberg extrapolation to eliminate the leading order correction Talay & Tubaro (1990).
- We carefully study the overdamped regime $\gamma \rightarrow +\infty$ in Section 2.6, making use in particular of uniform resolvent estimates obtained in Theorem 2.2 (based on a uniform hypocoercivity property on appropriate subspaces of $H^1(\mu)$);
- We provide error estimates for the computation of transport coefficients, by assessing the bias arising in the numerical discretization of either (i) the computation of integrated time-correlation functions expressing transport coefficients via Green-Kubo formulae; or (ii) ergodic averages of steady-state nonequilibrium dynamics where the equilibrium evolution (1.2) is perturbed by a non-gradient force and the transport coefficient is extracted from the linear response of some quantity of interest (see Section 3). The latter approach is illustrated by the study of the mobility, which measures the response in the average velocity arising from a constant external force exerted on the system. We also study the consistency of the numerical estimations in the overdamped limit.

Some numerical simulations are provided to illustrate the most important results (see Section 2.5.3 and 3.3).

2. Error estimates on the invariant measure for equilibrium dynamics

We start by giving some properties of the Langevin dynamics in Section 2.1 (most results are well-known, except for the material on the overdamped limit $\gamma \rightarrow +\infty$ presented in Section 2.1.3). The numerical schemes we consider are then described in Section 2.2, their ergodic properties being discussed in Section 2.3. Error estimates on the invariant measure are provided in Section 2.4. We then show in Section 2.5 how to estimate the leading order correction term through an appropriate integrated correlation function. An important side result of this section are error estimates for Green-Kubo type formulas. We finally study the errors on the invariant measures in the overdamped limit in Section 2.6.

2.1 Properties of equilibrium Langevin dynamics

Langevin dynamics can be seen as Hamiltonian dynamics perturbed by an Ornstein-Uhlenbeck process of magnitude $\gamma > 0$ in the momenta:

$$\begin{cases} dq_t = M^{-1} p_t dt, \\ dp_t = -\nabla V(q_t) dt - \gamma M^{-1} p_t dt + \sqrt{\frac{2\gamma}{\beta}} dW_t, \end{cases} \quad (2.1)$$

where W_t is a dN -dimensional standard Brownian motion, and M is the mass matrix of the system. We assume that $M = \text{diag}(m_1 I_d, \dots, m_N I_d)$, so that momenta are distributed according to independent Gaussian distributions under the canonical measure. Note that we formulate here the dynamics using friction forces proportional to the velocity of the particles.

The existence and uniqueness of strong solutions is guaranteed when the position space is compact since the kinetic energy function $1 + |p|^2$ is a Lyapunov function, see for instance (Rey-Bellet, 2006, Theorem 5.9). We will sometimes denote by $(q_{\gamma,t}, p_{\gamma,t})$ the solution of this equation to emphasize the dependence on the friction coefficient.

It is useful, to describe more conveniently splitting schemes, to introduce the elementary dynamics with generators

$$A = M^{-1} p \cdot \nabla_q, \quad B = -\nabla V(q) \cdot \nabla_p, \quad C = -M^{-1} p \cdot \nabla_p + \frac{1}{\beta} \Delta_p. \quad (2.2)$$

The generator \mathcal{L}_γ of the equilibrium Langevin dynamics (2.1) can then be written as

$$\mathcal{L}_\gamma = A + B + \gamma C,$$

where $\mathcal{L}_0 = A + B$ is the generator associated with the Hamiltonian part of the dynamics. The invariance of the canonical measure μ defined in (1.1) for the Langevin dynamics can be rewritten in terms of the generator \mathcal{L}_γ : for any smooth test function φ ,

$$\int_{\mathcal{E}} \mathcal{L}_\gamma \varphi d\mu = 0. \quad (2.3)$$

In fact, the operators $A + B$ and C separately preserve μ . In the sequel, we will by default consider $L^2(\mu)$ as the reference Hilbert space to define scalar products and associated norms, adjoints of operators, etc. In this functional setting, it is easy to check that

$$(A + B)^* = -(A + B), \quad C^* = C.$$

Note also that, thanks to a Poincaré inequality, the operator C has a compact resolvent on

$$L^2(\kappa) \cap \text{Ker}(\pi) = \left\{ f \in L^2(\kappa) \mid \int_{\mathbb{R}^{dN}} f(p) \kappa(dp) = 0 \right\},$$

with positive eigenvalues going to $+\infty$.

An important property of Langevin dynamics is its reversibility with respect to the invariant measure μ , up to momentum reversal (see the discussion in (Lelièvre *et al.*, 2010, Section 2.2.3)). In particular, introducing the unitary operator

$$(\mathcal{R}\varphi)(q, p) = \varphi(q, -p), \quad (2.4)$$

it holds that $\mathcal{R}\mathcal{L}_\gamma\mathcal{R} = \mathcal{L}_\gamma^*$ (with, as mentioned above, adjoints taken on $L^2(\mu)$).

2.1.1 Ergodicity results. The ergodicity of the Langevin dynamics for $\gamma > 0$, understood either as the almost sure convergence of time averages along a realization of the dynamics, or the long-time convergence of the law of the process to μ , is well established, see for instance Mattingly *et al.* (2002); Talay (2002); Cancès *et al.* (2007) and references therein. These references rely on the use of Lyapunov functions, following strategies of proofs pioneered in the Markov Chain community Meyn & Tweedie (2009), although alternative proofs relying on analytical tools exist Rey-Bellet (2006); Hairer & Mattingly (2011). In any case, the measure μ is the unique invariant measure of the dynamics. This property can be translated as $\text{Ker}(\mathcal{L}_\gamma) = \mathbb{C}\mathbf{1}$.

An alternative way to prove the long-time convergence of the law of the process is to use subelliptic or hypocoercive estimates Talay (2002); Eckmann & Hairer (2003); Hérau & Nier (2004); Villani (2009); Hairer & Pavliotis (2008). The interest of this approach is that it allows us to give more explicit rates of convergence than the ones obtained by Lyapunov type approaches, and gives fine results on the structure of the spectrum of \mathcal{L}_γ (see in particular Eckmann & Hairer (2003); Hérau & Nier (2004)). An important result of hypocoercivity in this case is that there exist $K_\gamma, \lambda_\gamma > 0$ such that

$$\|e^{t\mathcal{L}_\gamma}\|_{\mathcal{B}(\mathcal{H}^1)} \leq K_\gamma e^{-\lambda_\gamma t}, \quad (2.5)$$

where the subspace

$$\mathcal{H}^1 = H^1(\mu) \setminus \text{Ker}(\mathcal{L}_\gamma) = \left\{ u \in H^1(\mu) \mid \int_{\mathcal{E}} u d\mu = 0 \right\} \quad (2.6)$$

of the Hilbert space $H^1(\mu)$ is endowed with the norm $\|u\|_{H^1(\mu)}^2 = \|u\|_{L^2(\mu)}^2 + \|\nabla_p u\|_{L^2(\mu)}^2 + \|\nabla_q u\|_{L^2(\mu)}^2$, and $\|\cdot\|_{\mathcal{B}(\mathcal{H}^1)}$ is the operator norm on \mathcal{H}^1 . In particular, the operator \mathcal{L}_γ is invertible on \mathcal{H}^1 , and

$$\|\mathcal{L}_\gamma^{-1}\|_{\mathcal{B}(\mathcal{H}^1)} \leq \frac{K_\gamma}{\lambda_\gamma}.$$

Note that for unbounded position spaces, the potential V has to satisfy some assumptions for (2.5) to hold (such as a Poincaré inequality for $e^{-\beta V}$), but these assumptions are trivially satisfied when the position space is compact, as is the case here. An important issue is the dependence of the constant K_γ, λ_γ on γ , or at least the dependence of the resolvent norm $\|\mathcal{L}_\gamma^{-1}\|_{\mathcal{B}(\mathcal{H}^1)}$ on γ . This is made precise in the results presented below.

2.1.2 Hamiltonian limit $\gamma \rightarrow 0$. When $\gamma = 0$, the Langevin dynamics reduces to the Hamiltonian dynamics, whose generator $A + B$ has a kernel much larger than $\text{Ker}(\mathcal{L}_\gamma) = \mathbb{C}\mathbf{1}$. It is therefore expected that $\|\mathcal{L}_\gamma^{-1}\|_{\mathcal{B}(\mathcal{H}^1)}$ diverges as $\gamma \rightarrow 0$. The rate of divergence is made precise in the following theorem, summarizing the results from (Hairer & Pavliotis, 2008, Theorem 1.6 and Proposition 6.3).

THEOREM 2.1 (see Hairer & Pavliotis (2008)) Denote by $\|\cdot\|_{\mathcal{B}(\mathcal{H}^0)}$ the operator norm on the subspace

$$\mathcal{H}^0 = \left\{ u \in L^2(\mu) \mid \int_{\mathcal{E}} u d\mu = 0 \right\} \quad (2.7)$$

of the Hilbert space $L^2(\mu)$. There exists two constants $c_-, c_+ > 0$ such that, for any $0 < \gamma \leq 1$,

$$\frac{c_-}{\gamma} \leq \|\mathcal{L}_\gamma^{-1}\|_{\mathcal{B}(\mathcal{H}^0)} \leq \frac{c_+}{\gamma}.$$

2.1.3 Overdamped limit $\gamma \rightarrow +\infty$. The overdamped limit can be obtained by either letting the friction go to infinity in (2.1) together with an appropriate rescaling of time; or by letting masses go to 0. When discussing overdamped limits in this article, we will always set the mass matrix M to identity and consider the limit $\gamma \rightarrow +\infty$. Since we restrict our attention to the invariant measure of the system, the time rescaling is not relevant.

Let us describe more precisely the convergence result. It is shown in (Lelièvre *et al.*, 2010, Section 2.2.4) for instance that the solutions of (2.1) observed over long times, namely $(q_{\gamma, \gamma s}, p_{\gamma, \gamma s})_{s \geq 0}$, converge pathwise on finite time intervals $s \in [0, t]$ to the solutions of overdamped Langevin dynamics

$$dQ_t = -\nabla V(Q_t) dt + \sqrt{\frac{2}{\beta}} dW_t, \quad (2.8)$$

with the same initial condition $Q_0 = q_{\gamma, 0}$. The process (2.8) is ergodic on the compact position space \mathcal{M} , with unique invariant probability measure $\bar{\mu}(dq)$ defined in (1.4). Its generator

$$\mathcal{L}_{\text{ovd}} = -\nabla V(q) \cdot \nabla_q + \frac{1}{\beta} \Delta_q$$

is an elliptic operator which is self-adjoint on $L^2(\bar{\mu})$, with compact resolvent (see for instance the discussion and the references in (Lelièvre *et al.*, 2010, Section 2.3.2)). The inverse operator $\mathcal{L}_{\text{ovd}}^{-1}$ is bounded from

$$\tilde{H}^m(\bar{\mu}) = \left\{ \varphi \in H^m(\bar{\mu}) \mid \int_{\mathcal{M}} \varphi d\bar{\mu} = 0 \right\}$$

to $\tilde{H}^{m+2}(\bar{\mu})$.

The following result gives bounds on the resolvent of the Langevin generator in the overdamped regime, and in fact quantifies the difference between the resolvent $\mathcal{L}_{\gamma}^{-1}$ and the resolvent $\mathcal{L}_{\text{ovd}}^{-1}$ appropriately rescaled by a factor γ .

THEOREM 2.2 There exist two constants $c_-, c_+ > 0$ such that, for any $\gamma \geq 1$,

$$c_- \gamma \leq \|\mathcal{L}_{\gamma}^{-1}\|_{\mathcal{B}(\mathcal{H}^1)} \leq c_+ \gamma. \quad (2.9)$$

More precisely, there exists a constant $K > 0$ such that, for any $\gamma \geq 1$,

$$\begin{aligned} \left\| \mathcal{L}_{\gamma}^{-1} - \gamma \mathcal{L}_{\text{ovd}}^{-1} \pi - p^T \nabla_q \mathcal{L}_{\text{ovd}}^{-1} \pi + \mathcal{L}_{\text{ovd}}^{-1} \pi (A+B) C^{-1} (\text{Id} - \pi) \right\|_{\mathcal{B}(\mathcal{H}^1)} &\leq \frac{K}{\gamma}, \\ \left\| (\mathcal{L}_{\gamma}^*)^{-1} - \gamma \mathcal{L}_{\text{ovd}}^{-1} \pi + p^T \nabla_q \mathcal{L}_{\text{ovd}}^{-1} \pi - \mathcal{L}_{\text{ovd}}^{-1} \pi (A+B) C^{-1} (\text{Id} - \pi) \right\|_{\mathcal{B}(\mathcal{H}^1)} &\leq \frac{K}{\gamma}, \end{aligned} \quad (2.10)$$

where the operator π is defined in (1.5), and $(C^{-1}\psi)(q, p)$ is understood as applying the operator C^{-1} to the function $\psi(q, \cdot) \in L^2(\kappa)$ for all values of $q \in \mathcal{M}$.

Note that the function $\mathcal{L}_{\text{ovd}}^{-1} \pi f$ is well defined since, as f belongs to \mathcal{H}^1 , the function πf has a vanishing average with respect to $\bar{\mu}$. An important ingredient in the proof of Theorem 2.2 is the following estimate, which we call uniform hypocoercivity estimate.

LEMMA 2.1 (Uniform hypocoercivity for large frictions) Consider the following subspace of \mathcal{H}^1 :

$$\mathcal{H}_{\perp}^1 = \left\{ u \in \mathcal{H}^1 \mid \bar{u}(q) = \int_{\mathbb{R}^{dN}} u(q, p) \kappa(dp) \text{ is constant} \right\}.$$

There exists a constant $K > 0$ such that, for any $\gamma \geq 1$, it holds

$$\forall f \in \mathcal{H}_{\perp}^1, \quad \|\mathcal{L}_{\gamma}^{-1} f\|_{H^1(\mu)} \leq K \|f\|_{H^1(\mu)}.$$

The proofs of Theorem 2.2 and Lemma 2.1 are provided in Section 4.1.

2.2 Splitting schemes for the equilibrium Langevin dynamics

We present in this section the splitting schemes to be examined in this article. In fact, these schemes are described in terms of evolution operators $P_{\Delta t}$ which are such that the Markov chain (q^n, p^n) generated by the discretization satisfies

$$P_{\Delta t} \psi(q, p) = \mathbb{E} \left(\psi(q^{n+1}, p^{n+1}) \mid (q^n, p^n) = (q, p) \right).$$

We also briefly give some ergodicity results obtained by minor extensions or variations of existing results in the literature (see in particular Mattingly *et al.* (2002); Talay (2002); Bou-Rabee & Owhadi (2010)). Since these ergodicity issues are by now a rather standard and well-understood matter, especially for compact position spaces, we provide only elements of proofs in Section 4.2.

2.2.1 First-order splitting schemes. First-order schemes are obtained by a Lie-Trotter splitting of the elementary evolutions generated by $A, B, \gamma C$. The motivation for this splitting is that all elementary evolutions are analytically integrable (see the expressions of the associated semigroups in (4.14)). There are 6 possible schemes, whose evolution operators are of the general form

$$P_{\Delta t}^{ZYX} = e^{\Delta t Z} e^{\Delta t Y} e^{\Delta t X},$$

with all possible permutations (Z, Y, X) of $(A, B, \gamma C)$. For instance, the numerical scheme associated with $P_{\Delta t}^{B,A,\gamma C}$ is

$$\begin{cases} \tilde{p}^{n+1} = p^n - \Delta t \nabla V(q^n), \\ q^{n+1} = q^n + \Delta t M^{-1} \tilde{p}^{n+1}, \\ p^{n+1} = \alpha_{\Delta t} \tilde{p}^{n+1} + \sqrt{\frac{1 - \alpha_{\Delta t}^2}{\beta}} M G^n, \end{cases} \quad (2.11)$$

where $\alpha_{\Delta t} = \exp(-\gamma M^{-1} \Delta t)$, and G^n are independent and identically distributed standard Gaussian random variables. Note that the order of the operations performed on the configuration of the system is the inverse of the order of the operations mentioned in the superscript of the evolution operator $P_{\Delta t}^{B,A,\gamma C}$ when read from right to left. This inversion is known as *Vertauschungssatz* (see for instance the discussion in (Hairer *et al.*, 2006, Section III.5.1)). It arises from the fact that the numerical method modifies the distribution of the variables, whereas the evolution operator encodes the evolution of observables (determined by the adjoint of the operator encoding the evolution of the distribution).

The iterations of the three schemes associated with $P_{\Delta t}^{\gamma C,B,A}$, $P_{\Delta t}^{B,A,\gamma C}$, $P_{\Delta t}^{A,\gamma C,B}$ share a common sequence of update operations, as for $P_{\Delta t}^{\gamma C,A,B}$, $P_{\Delta t}^{A,B,\gamma C}$, $P_{\Delta t}^{B,\gamma C,A}$. More precisely, we mean that equalities of the following form hold:

$$\left(P_{\Delta t}^{A,B,\gamma C}\right)^n = T_{\gamma,\Delta t} \left(P_{\Delta t}^{\gamma C,A,B}\right)^{n-1} U_{\Delta t}, \quad U_{\Delta t} = e^{\gamma \Delta t C}, \quad T_{\gamma,\Delta t} = e^{\Delta t A} e^{\Delta t B}. \quad (2.12)$$

It is therefore not surprising that the invariant measures of the schemes with operators composed in the same order have very similar properties, as made precise in Theorem 2.5, relying on Lemma 2.4.

2.2.2 Second-order schemes. Second-order schemes are obtained by a Strang splitting of the elementary evolutions generated by $A, B, \gamma C$. There are also 6 possible schemes, which are of the general form

$$P_{\Delta t}^{ZYYYYZ} = e^{\Delta t Z/2} e^{\Delta t Y/2} e^{\Delta t X} e^{\Delta t Y/2} e^{\Delta t Z/2},$$

with the same possible orderings as for first-order schemes. Again, these schemes can be classified into 3 groups depending on the ordering of the operators once the elementary one-step evolution is iterated: (i) $P_{\Delta t}^{\gamma C,B,A,B,\gamma C}$, $P_{\Delta t}^{A,B,\gamma C,B,A}$, (ii) $P_{\Delta t}^{\gamma C,A,B,A,\gamma C}$, $P_{\Delta t}^{B,A,\gamma C,A,B}$, and (iii) $P_{\Delta t}^{B,\gamma C,A,\gamma C,B}$, $P_{\Delta t}^{A,\gamma C,B,\gamma C,A}$. We discard the latter category since the invariant measures of the associated numerical schemes are not consistent with $\bar{\mu}$ in the overdamped limit (see Section 2.6).

2.2.3 Geometric Langevin Algorithms. In fact, as already proved in Bou-Rabee & Owhadi (2010) (see also Corollary 2.2 below), a second-order accuracy on the invariant measure can be obtained by resorting to a first-order splitting between the Hamiltonian and the Ornstein-Uhlenbeck parts, and discretizing the Hamiltonian part with a second-order scheme. This corresponds to the following evolution operators of Geometric Langevin Algorithm (GLA) type:

$$\begin{aligned} P_{\Delta t}^{\gamma C,A,B,A} &= e^{\gamma \Delta t C} e^{\Delta t A/2} e^{\Delta t B} e^{\Delta t A/2}, & P_{\Delta t}^{\gamma C,B,A,B} &= e^{\gamma \Delta t C} e^{\Delta t B/2} e^{\Delta t A} e^{\Delta t B/2}, \\ P_{\Delta t}^{A,B,A,\gamma C} &= e^{\Delta t A/2} e^{\Delta t B} e^{\Delta t A/2} e^{\gamma \Delta t C}, & P_{\Delta t}^{B,A,B,\gamma C} &= e^{\Delta t B/2} e^{\Delta t A} e^{\Delta t B/2} e^{\gamma \Delta t C}. \end{aligned} \quad (2.13)$$

2.3 Ergodicity results for splitting schemes

Let us now give some technical results on the ergodic behavior of the splitting schemes presented in Section 2.2, generically denoting in this section by $P_{\Delta t}$ the evolution operator (we do not denote explicitly the dependence on the friction parameter γ although the constants appearing in the results below a priori depend on this parameter). Ergodicity results for a fixed value of Δt are obtained with techniques similar to the ones presented in Meyn & Tweedie (2009), by mimicking the proofs presented for certain discretization schemes of the Langevin equation in Mattingly *et al.* (2002); Talay (2002); Bou-Rabee & Owhadi (2010). A more subtle point is to obtain rates of convergence which are uniform in the time-step Δt . We are able here to prove such results by relying on the fact that the position space \mathcal{M} is compact.

The proof is based on two preliminary results, namely a uniform drift inequality or Lyapunov condition and a uniform minorization condition (see Section 4.2 for the proof). The term uniform refers to estimates which are independent of the time-step Δt . To obtain such estimates, we have to consider evolutions over fixed times $T \simeq n\Delta t$, which amounts to iterating the elementary evolution $P_{\Delta t}$ over $\lceil T/\Delta t \rceil$ time steps (where $\lceil x \rceil$ denotes the smallest integer larger than x).

LEMMA 2.2 (Uniform Lyapunov condition) Consider the family of Lyapunov functions for $s \in \mathbb{N}^*$:

$$\mathcal{K}_s(q, p) = 1 + |p|^{2s}. \quad (2.14)$$

For any $s^* \in \mathbb{N}^*$, there exist $\Delta t^* > 0$ and $C_a, C_b > 0$ such that, for any $1 \leq s \leq s^*$ and $0 < \Delta t \leq \Delta t^*$,

$$P_{\Delta t} \mathcal{K}_s \leq a_{\Delta t} \mathcal{K}_s + b_{\Delta t}, \quad 0 \leq a_{\Delta t} \leq \exp(-C_a \Delta t), \quad 0 \leq b_{\Delta t} \leq C_b \Delta t. \quad (2.15)$$

In particular, for any $T > 0$,

$$P_{\Delta t}^{\lceil T/\Delta t \rceil} \mathcal{K}_s \leq \exp(-C_a T) \mathcal{K}_s + \frac{2C_b}{C_a}. \quad (2.16)$$

LEMMA 2.3 (Uniform minorization condition) Consider $T > 0$ sufficiently large, and fix any $p_{\max} > 0$. There exist $\Delta t^*, \alpha > 0$ and a probability measure ν such that, for any bounded, measurable non-negative function f , and any $0 < \Delta t \leq \Delta t^*$,

$$\inf_{|p| \leq p_{\max}} \left(P_{\Delta t}^{\lceil T/\Delta t \rceil} f \right)(q, p) \geq \alpha \int_{\mathcal{E}} f(q, p) \nu(dq dp).$$

Lemma 2.3 ensures that the Assumption 2 in Hairer & Mattingly (2011) holds for any choice of Lyapunov function \mathcal{K}_s ($s \geq 1$), provided p_{\max} is chosen to be sufficiently large. The uniform minorization condition can formally be rewritten as

$$\forall (q_0, p_0) \in \mathcal{M} \times B(0, p_{\max}), \quad P_{\Delta t}((q_0, p_0), dq dp) \geq \alpha \nu(dq dp).$$

We present a direct proof of Lemma 2.3 in Section 4.2. Extending this result to unbounded position spaces is much more difficult in general, see for instance recent works assuming non-degeneracy of the noise Klokov & Veretennikov (2006, 2013); Bou-Rabee & Hairer (2013).

Let us now precisely state the ergodicity result. To this end, we need to define the Banach spaces

$$L_{\mathcal{K}_s}^\infty = \left\{ \psi \text{ measurable} \mid \frac{\psi}{\mathcal{K}_s} \in L^\infty(\mathcal{E}) \right\},$$

endowed with the norms

$$\|\psi\|_{L_{\mathcal{K}_s}^\infty} = \left\| \frac{\psi}{\mathcal{K}_s} \right\|_{L^\infty}.$$

PROPOSITION 2.3 (Ergodicity of numerical schemes) Fix $s^* \geq 1$. For any $0 < \gamma < +\infty$, there exists Δt^* such that, for any $0 < \Delta t \leq \Delta t^*$, the Markov chain associated with $P_{\Delta t}$ has a unique invariant probability measure $\mu_{\gamma, \Delta t}$, which admits a density with respect to the Lebesgue measure $dq dp$, and has finite moments: For any $1 \leq s \leq s^*$,

$$\int_{\mathcal{E}} \mathcal{K}_s d\mu_{\gamma, \Delta t} \leq \frac{b_{\Delta t}}{1 - a_{\Delta t}} \leq \frac{C_b}{C_a} < +\infty, \quad (2.17)$$

uniformly in the time step Δt . There also exist $\lambda, K > 0$ (depending on s^* and γ but not on Δt) such that, for almost all $(q, p) \in \mathcal{E}$, and for all functions $f \in L_{\mathcal{K}_s}^\infty$,

$$\forall n \in \mathbb{N}, \quad \left| (P_{\Delta t}^n f)(q, p) - \int_{\mathcal{E}} f d\mu_{\gamma, \Delta t} \right| \leq K \mathcal{K}_s(q, p) e^{-\lambda n \Delta t} \|f\|_{L_{\mathcal{K}_s}^\infty}. \quad (2.18)$$

Let us emphasize again that, compared to the results of Mattingly *et al.* (2002); Talay (2002); Bou-Rabee & Owhadi (2010), the only new estimate is the uniform-in- Δt decay rate in (2.18), which follows from an application of the results of Hairer & Mattingly (2011) to the sampled chain $P_{\Delta t}^{\lceil T/\Delta t \rceil}$ (see Section 4.2 for further precisions). Recall also that the convergence rates we obtain depend on the friction parameter γ .

An interesting consequence of the above estimates is that we are able to obtain a uniform control on the resolvent of the operator $\text{Id} - P_{\Delta t}$. Such a bound will prove useful to control approximation errors in Green-Kubo type formulas (see Section 2.5). Note indeed that the estimate (2.18) implies the operator bound

$$\|P_{\Delta t}^n\|_{\mathcal{B}(L_{\mathcal{K}_s, \Delta t}^\infty)} \leq K e^{-\lambda n \Delta t},$$

on the Banach space

$$L_{\mathcal{K}_s, \Delta t}^\infty = \left\{ \psi \in L_{\mathcal{K}_s}^\infty \mid \int_{\mathcal{E}} \psi d\mu_{\gamma, \Delta t} = 0 \right\}.$$

Note that $L_{\mathcal{K}_s, \Delta t}^\infty$ depends both on Δt and γ through $\mu_{\gamma, \Delta t}$, although the dependence on γ is not explicitly written. This proves that the series

$$\sum_{n=0}^{+\infty} P_{\Delta t}^n$$

is well defined as a bounded operator on $L_{\mathcal{K}_s, \Delta t}^\infty$, and in fact is equal to $(\text{Id} - P_{\Delta t})^{-1}$ since

$$(\text{Id} - P_{\Delta t}) \sum_{n=0}^{+\infty} P_{\Delta t}^n = \text{Id}.$$

We also have the bound

$$\left\| (\text{Id} - P_{\Delta t})^{-1} \right\|_{\mathcal{B}(L_{\mathcal{K}_s, \Delta t}^\infty)} \leq \sum_{n=0}^{+\infty} \|P_{\Delta t}^n\|_{\mathcal{B}(L_{\mathcal{K}_s, \Delta t}^\infty)} \leq \frac{K}{1 - e^{-\lambda \Delta t}} \leq \frac{2K}{\lambda \Delta t}$$

provided Δt is sufficiently small. Let us summarize this result as follows.

COROLLARY 2.1 For any $s^* \in \mathbb{N}^*$, there exist $\Delta t^* > 0$ and $R > 0$ such that, for all $0 \leq s \leq s^*$, a uniform resolvent bound holds: for any $0 < \Delta t \leq \Delta t^*$,

$$\left\| \left(\frac{\text{Id} - P_{\Delta t}}{\Delta t} \right)^{-1} \right\|_{\mathcal{B}(L_{\mathcal{K}_s, \Delta t}^\infty)} \leq R. \quad (2.19)$$

2.4 Error estimates for finite frictions

In this section we study the error on the average of sufficiently smooth functions, which allows us to characterize the corrections to the invariant measure. In Theorems 2.5 and 2.7, below, we characterize all the first- and second-order splittings; the technique of proof allows us to provide a rigorous study of the error estimates in the overdamped regime (see Section 2.6) and for nonequilibrium systems (see Section 3). If only the order of magnitude of the correction is of interest, and not the expression of the correction in itself, no regularity on the derivatives is required (see Bou-Rabee & Owhadi (2010)), in contrast to situations where such corrections are explicitly considered, as in Talay (2002) for instance. In the latter article, results are stated for smooth functions having a sufficient number of derivatives growing at most

polynomially. Here and in the sequel, we will also consider such functions. Note that, since the position space is compact, only the growth in the momentum variable has to be controlled.

For some function $\mathcal{K} \geq 1$, let us introduce the spaces $W_{\mathcal{K}}^{n,\infty}$ defined recursively as $W_{\mathcal{K}}^{0,\infty} = L_{\mathcal{K}}^{\infty}$ and

$$W_{\mathcal{K}}^{n,\infty} = \left\{ f \in L_{\mathcal{K}}^{\infty} \mid \nabla f \in \left(W_{\mathcal{K}}^{n-1,\infty} \right)^{2dN} \right\}.$$

Note that $W_{\mathcal{K}}^{n,\infty} \subset H^n(\mu)$ when the function \mathcal{K} is in $L^2(\mu)$ (since in this case $L_{\mathcal{K}}^{\infty} \subset L^2(\mu)$).

DEFINITION 2.4 (Sufficiently smooth functions) The set \mathcal{S} of smooth functions is the set of functions $f \in L^2(\mu)$ such that, for any $m \geq 0$, there exists $s \geq 0$ (depending on f and m) so that $f \in W_{\mathcal{K}_s}^{m,\infty}$ (with \mathcal{K}_s defined in (2.14)).

A slight extension of the results of Talay (2002) (briefly sketched in Section 4.3.1) allows us to show that the set

$$\widetilde{\mathcal{S}} = \left\{ f \in \mathcal{S} \mid \int_{\mathcal{E}} f d\mu = 0 \right\}$$

is stable with respect to \mathcal{L}^{-1} . This result could also probably be obtained by appropriately modifying the proofs from Eckmann & Hairer (2003); Hérau & Nier (2004) to account for compact position spaces. Let us finally mention that the set $\mathcal{S} \cap \text{Ker}(\pi)$ is of course stable with respect to $\mathcal{L}_{\text{ovd}}^{-1}$.

2.4.1 Relating invariant measures of two numerical schemes. We classified in Section 2.2 the numerical schemes according to the order of appearance of the elementary operators. More precisely, we considered schemes to be similar when the global ordering of the operators is the same but the operations are started and ended differently, as in (2.12) above (see also (2.20) below for an abstract definition). We motivate here why we resorted to this classification: It is indeed straightforward to obtain the expression of the invariant measure of one scheme when the expression for another one is given.

We state the result in an abstract fashion for two schemes $P_{\Delta t} = U_{\Delta t} T_{\Delta t}$ and $Q_{\Delta t} = T_{\Delta t} U_{\Delta t}$ (which implies the condition (2.20) below). See (2.12) for a concrete example.

LEMMA 2.4 (Here and elsewhere: TU lemma) Consider two numerical schemes with associated evolution operators $P_{\Delta t}, Q_{\Delta t}$, and for which there exist operators $U_{\Delta t}, T_{\Delta t}$ such that, for all $n \geq 1$,

$$Q_{\Delta t}^n = T_{\Delta t} P_{\Delta t}^{n-1} U_{\Delta t}. \quad (2.20)$$

We also assume that both schemes are ergodic with associated invariant measures denoted respectively by $\mu_{P,\Delta t}, \mu_{Q,\Delta t}$: For almost all $(q, p) \in \mathcal{E}$ and all bounded measurable functions f ,

$$\lim_{n \rightarrow +\infty} P_{\Delta t}^n f(q, p) = \int_{\mathcal{E}} f d\mu_{P,\Delta t}, \quad \lim_{n \rightarrow +\infty} Q_{\Delta t}^n f(q, p) = \int_{\mathcal{E}} f d\mu_{Q,\Delta t}.$$

Then, for all bounded measurable functions φ ,

$$\int_{\mathcal{E}} \varphi d\mu_{Q,\Delta t} = \int_{\mathcal{E}} (U_{\Delta t} \varphi) d\mu_{P,\Delta t}. \quad (2.21)$$

The proof of this result relies on the simple observation that, for a given initial measure ρ with a smooth density with respect to the Lebesgue measure, the ergodicity assumption (implied by conditions such as (2.18)) ensures that

$$\int_{\mathcal{E}} \varphi d\mu_{Q,\Delta t} = \lim_{n \rightarrow +\infty} \int_{\mathcal{E}} Q_{\Delta t}^n \varphi d\rho = \lim_{n \rightarrow +\infty} \int_{\mathcal{E}} T_{\Delta t} P_{\Delta t}^{n-1} (U_{\Delta t} \varphi) d\rho = \int_{\mathcal{E}} (U_{\Delta t} \varphi) d\mu_{P,\Delta t}.$$

Let us now show how we will use Lemma 2.4 in the sequel. Assume that a weak error estimate holds on the invariant measure $\mu_{P,\Delta t}$: there exist $\alpha \geq 1$ and a function $f_{\alpha} \in \mathcal{S}$ such that

$$\int_{\mathcal{E}} \psi d\mu_{P,\Delta t} = \int_{\mathcal{E}} \psi d\mu + \Delta t^{\alpha} \int_{\mathcal{E}} \psi f_{\alpha} d\mu + \Delta t^{\alpha+1} r_{\psi,\alpha,\Delta t},$$

with $|r_{\psi,\alpha,\Delta t}| \leq K$ for Δt sufficiently small. Combining this equality and (2.21), the following expansion is obtained for $\mu_{Q,\Delta t}$:

$$\int_{\mathcal{E}} \psi d\mu_{Q,\Delta t} = \int_{\mathcal{E}} (U_{\Delta t} \psi) d\mu_{P,\Delta t} = \int_{\mathcal{E}} (U_{\Delta t} \psi) d\mu + \Delta t^\alpha \int_{\mathcal{E}} (U_{\Delta t} \psi) f_\alpha d\mu + \Delta t^{\alpha+1} r_{U_{\Delta t} \psi, \alpha, \Delta t}.$$

In general, for an evolution operator $U_{\Delta t}$ preserving the measure μ at order $\delta \geq 1$, we can write

$$U_{\Delta t} = \text{Id} + \Delta t \mathcal{A}_1 + \cdots + \Delta t^{\delta-1} \mathcal{A}_{\delta-1} + \Delta t^\delta S_\delta + \Delta t^{\delta+1} R_{\delta,\Delta t},$$

where the operators \mathcal{A}_k preserve the measure μ (equalities such as (2.3) hold with \mathcal{L} replaced by \mathcal{A}_k), and the operator S_δ does not. Typically, \mathcal{A}_k is a composition of the operators $A + B$ and C . Three cases should then be distinguished:

- (i) When $\delta \geq \alpha + 1$, the weak error in the invariant measure $\mu_{Q,\Delta t}$ is of the same order as for $\mu_{P,\Delta t}$ since

$$\int_{\mathcal{E}} \psi d\mu_Q = \int_{\mathcal{E}} \psi d\mu + \Delta t^\alpha \int_{\mathcal{E}} \psi f_\alpha d\mu + \Delta t^{\alpha+1} \tilde{r}_{\psi, \alpha, \delta, \Delta t}.$$

- (ii) For $\delta \leq \alpha - 1$, the weak error in the invariant measure μ_Q arises at dominant order from the operator $U_{\Delta t}$:

$$\int_{\mathcal{E}} \psi d\mu_Q = \int_{\mathcal{E}} \psi d\mu + \Delta t^\delta \int_{\mathcal{E}} \psi (S_\delta^* \mathbf{1}) d\mu + \Delta t^{\delta+1} \tilde{r}_{\psi, \alpha, \delta, \Delta t}.$$

- (iii) The interesting case corresponds to $\alpha = \delta$. In this situation,

$$\int_{\mathcal{E}} \psi d\mu_Q = \int_{\mathcal{E}} \psi d\mu + \Delta t^\alpha \int_{\mathcal{E}} \psi (f_\alpha + S_\alpha^* \mathbf{1}) d\mu + \Delta t^{\alpha+1} \tilde{r}_{\psi, \alpha, \delta, \Delta t}. \quad (2.22)$$

An increase in the order of the error on the invariant measure is obtained when the leading order correction vanishes for all admissible observables ψ , that is, if and only if $f_\alpha + S_\alpha^* \mathbf{1} = 0$.

2.4.2 First-order schemes. The following result characterizes at leading order the invariant measure of the schemes based on a first-order splitting (see Section 2.2.1). We first study the error estimates in the invariant measure of the schemes $P_{\Delta t}^{\gamma C, B, A}$, $P_{\Delta t}^{\gamma C, A, B}$ (which can be interpreted as GLA schemes with a symplectic Euler discretization of the Hamiltonian part Bou-Rabee & Owhadi (2010)), and then deduce error estimates for the four remaining schemes introduced in Section 2.2.1 by making use of Lemma 2.4. The proof can be found in Section 4.4.

THEOREM 2.5 Consider any of the first order splittings presented in Section 2.2.1, and denote by $\mu_{\gamma, \Delta t}(dq dp)$ its invariant measure. Then there exists a smooth function $f_{1,\gamma}$ such that, for any function $\psi \in \mathcal{S}$,

$$\int_{\mathcal{E}} \psi(q, p) \mu_{\gamma, \Delta t}(dq dp) = \int_{\mathcal{E}} \psi(q, p) \mu(dq dp) + \Delta t \int_{\mathcal{E}} \psi(q, p) f_{1,\gamma}(q, p) \mu(dq dp) + \Delta t^2 r_{\psi, \gamma, \Delta t}, \quad (2.23)$$

where the remainder $r_{\psi, \gamma, \Delta t}$ is uniformly bounded for Δt sufficiently small. The expressions of the correction functions $f_{1,\gamma}$ depend on the numerical scheme at hand. They are defined as

$$\begin{aligned} \mathcal{L}_\gamma^* f_1^{\gamma C, B, A} &= -\frac{1}{2}(A+B)g, & g(q, p) &= \beta p^T M^{-1} \nabla V(q), \\ f_1^{\gamma C, A, B} &= f_1^{A, B, \gamma C} = -f_1^{B, A, \gamma C} = -f_1^{\gamma C, B, A}, \\ f_1^{A, \gamma C, B} &= -f_1^{B, \gamma C, A} = f_1^{\gamma C, B, A} - g. \end{aligned} \quad (2.24)$$

It is in fact possible to uniformly control the remainder $r_{\psi, \gamma, \Delta t}$ thanks to functional inequalities such as (4.12).

REMARK 2.1 Note that the equations (2.24) could be analytically solved if, instead of the fluctuation/dissipation operator C , we were using the mass-weighted differential operator as in Leimkuhler & Matthews (2013a):

$$C_M = -p^T \nabla_p + \frac{1}{\beta} M : \nabla_p^2.$$

The corresponding generator $\mathcal{L}_{\gamma,M} = A + B + \gamma C_M$ is associated with a Langevin dynamics where the friction force is proportional to the momenta rather than velocities. A simple computation shows that

$$-\frac{1}{2}(A+B)g = \mathcal{L}_{\gamma,M}^* \left(\frac{\beta}{2} V - g \right).$$

The condition (2.24) would be replaced by $\mathcal{L}_{\gamma,M}^* f_1^{\gamma C,B,A} = -(A+B)g/2$, so that $f_1^{\gamma C,B,A} = \beta V/2 - g + c$ where c is a constant ensuring that $f_1^{\gamma C,B,A}$ has a vanishing average with respect to μ .

2.4.3 Hamiltonian limit of the correction term. For first order splitting schemes, the limit of the leading order correction term in (2.23) can be studied in the limit when $\gamma \rightarrow 0$. Not surprisingly, it turns out that the leading order correction is the first term in the expansion of the modified Hamiltonian in powers of Δt . In contrast to the more complete proof we are able to present for the overdamped limit (see Section 2.6), we were not able to study the behavior of the remainder terms $r_{\psi,\gamma,\Delta t}$ in (2.23). There is a technical obstruction to controlling these remainders from the way we prove our results since the limiting operator $\mathcal{L}_0 = A + B$ is not invertible. Let us also mention that studying the corresponding Hamiltonian limit for second order schemes turns out to be a much more difficult question (see Remark 2.2).

PROPOSITION 2.6 There exists a constant $K > 0$ such that, for all $0 < \gamma \leq 1$,

$$\left\| f_1^{\gamma C,B,A} - \frac{\beta}{2} p^T M^{-1} \nabla V \right\|_{L^2(\mu)} \leq K\gamma,$$

with similar estimates for $f_1^{B,\gamma C,A}$ and $f_1^{B,A,\gamma C}$; and

$$\left\| f_1^{\gamma C,A,B} + \frac{\beta}{2} p^T M^{-1} \nabla V \right\|_{L^2(\mu)} \leq K\gamma,$$

with similar estimates for $f_1^{A,\gamma C,B}$ and $f_1^{A,B,\gamma C}$.

The proof of this result is provided in Section 4.5.

2.4.4 Second-order schemes. The following result characterizes at leading order the invariant measure of the schemes based on a second-order splitting (see Section 2.2.2).

THEOREM 2.7 Consider any of the second order splittings presented in Section 2.2.2, and denote by $\mu_{\gamma,\Delta t}(dq dp)$ its invariant measure. Then there exists a smooth function $f_{2,\gamma}$ such that, for any function $\psi \in \mathcal{S}$,

$$\int_{\mathcal{E}} \psi(q, p) \mu_{\gamma,\Delta t}(dq dp) = \int_{\mathcal{E}} \psi(q, p) \mu(dq dp) + \Delta t^2 \int_{\mathcal{E}} \psi(q, p) f_{2,\gamma}(q, p) \mu(dq dp) + \Delta t^4 r_{\psi,\gamma,\Delta t}, \quad (2.25)$$

where the remainder $r_{\psi,\gamma,\Delta t}$ is uniformly bounded for Δt sufficiently small. The expressions of the correction functions $f_{2,\gamma}$ depend on the numerical scheme at hand. They are defined as

$$\begin{aligned} \mathcal{L}_{\gamma}^* f_2^{\gamma C,B,A,B,\gamma C} &= \frac{1}{12}(A+B) \left[\left(A + \frac{B}{2} \right) g \right], & g(q, p) &= \beta p^T M^{-1} \nabla V(q), \\ \mathcal{L}_{\gamma}^* f_2^{\gamma C,A,B,A,\gamma C} &= -\frac{1}{12}(A+B) \left[\left(B + \frac{A}{2} \right) g \right], \\ f_2^{A,B,\gamma C,B,A} &= f_2^{\gamma C,B,A,B,\gamma C} + \frac{1}{8}(A+B)g, \\ f_2^{B,A,\gamma C,A,B} &= f_2^{\gamma C,A,B,A,\gamma C} - \frac{1}{8}(A+B)g. \end{aligned} \quad (2.26)$$

It can be checked that the expressions of $f_2^{B,A,\gamma C,A,B}$ and $f_2^{A,B,\gamma C,B,A}$ agree with the ones presented in Leimkuhler & Matthews (2013a). Let us emphasize that no Δt^3 correction term appears in (2.25) after the Δt^2 term. In fact, a more careful treatment would allow us to write an error expansion in terms of higher orders of Δt , with only even powers of Δt appearing.

The proof of this result is given in Section 4.6. We use as reference schemes for the proofs the schemes $p_{\Delta t}^{\gamma C,A,B,A,\gamma C}$, $p_{\Delta t}^{\gamma C,B,A,B,\gamma C}$. These schemes indeed turn out to be particularly convenient to study the over-damped limit.

The results from Theorem 2.7 allow us to obtain error estimates for the so-called Geometric Langevin Algorithms (GLA) introduced in Bou-Rabee & Owahdi (2010). Recall the somewhat surprising result that the error in the invariant measure of the GLA schemes is of order Δt^p for a discretization of order p of the Hamiltonian part, even though the strong order of the scheme is only one. The following result complements the estimate given in Bou-Rabee & Owahdi (2010) by making precise the leading order corrections to the invariant measure of the numerical scheme with respect to the canonical measure (see the proof in Section 4.7).

COROLLARY 2.2 (Error estimates for GLA schemes) Consider one of the GLA schemes defined in (2.13), and denote by $\mu_{\gamma,\Delta t}(dq dp)$ its invariant measure. Then there exist smooth functions $f_{2,\gamma}$ and $f_{3,\gamma}$ such that, for any function $\psi \in \mathcal{S}$,

$$\begin{aligned} \int_{\mathcal{E}} \psi(q, p) \mu_{\gamma,\Delta t}(dq dp) &= \int_{\mathcal{E}} \psi(q, p) \mu(dq dp) + \Delta t^2 \int_{\mathcal{E}} \psi(q, p) f_{2,\gamma}(q, p) \mu(dq dp) \\ &\quad + \Delta t^3 \int_{\mathcal{E}} \psi(q, p) f_{3,\gamma}(q, p) \mu(dq dp) + \Delta t^4 r_{\psi,\gamma,\Delta t}, \end{aligned} \quad (2.27)$$

where the remainder $r_{\psi,\gamma,\Delta t}$ is uniformly bounded for Δt sufficiently small. The expressions of the correction functions $f_{2,\gamma}$ and $f_{3,\gamma}$ are

$$\begin{aligned} f_2^{\gamma C,A,B,A} &= f_2^{\gamma C,A,B,A,\gamma C}, & f_3^{\gamma C,A,B,A} &= -\frac{\gamma}{2} C f_2^{\gamma C,A,B,A}, \\ f_2^{\gamma C,B,A,B} &= f_2^{\gamma C,B,A,B,\gamma C}, & f_3^{\gamma C,B,A,B} &= -\frac{\gamma}{2} C f_2^{\gamma C,B,A,B}. \end{aligned} \quad (2.28)$$

Note that the leading order term of the error is the same as for the corresponding second order splitting schemes. The next order correction (of order Δt^3) vanishes for functions ψ depending only on the position variable q .

REMARK 2.2 (Hamiltonian limit of the correction functions $f_{2,\gamma}$) Proving a result similar to Proposition 2.6 for second order splitting schemes or GLA schemes turns out to be much more difficult, although we formally expect that the limit of $f_{2,\gamma}$ as $\gamma \rightarrow 0$ is the first order correction of the modified Hamiltonian constructed by backward analysis. From (2.26), it should indeed be the case that $f_2^{\gamma C,B,A,B,\gamma C}$ converges to

$$f_2^{B,A,B} = -\frac{1}{12} \left(A + \frac{B}{2} \right) g.$$

Moreover, as we already mentioned before Proposition 2.6, we are not able to uniformly control remainder terms in the error expansion (2.25) as $\gamma \rightarrow 0$.

2.5 Numerical estimation of the correction term

The results of Section 2.4 show that the leading order correction terms for an observable ψ can be written as

$$\int_{\mathcal{E}} \psi(q, p) f_{\gamma}(q, p) \mu(dq dp), \quad (2.29)$$

where the function f_{γ} is the solution of a Poisson equation

$$\mathcal{L}_{\gamma}^* f_{\gamma} = g_{\gamma}, \quad (2.30)$$

the function g_γ depending on the numerical scheme at hand. It is in general impossible to analytically solve (2.30), and very difficult to numerically approximate its solution since it is a very high-dimensional partial differential equation. It is however possible to rewrite (2.29) as some integrated correlation function, a quantity which is amenable to numerical approximation. This is a standard way of computing transport coefficients based on Green-Kubo formulae, see the summary provided in Section 3.1. It provides here a way to compute the first order correction in the perfect sampling bias with a single simulation (as an alternative to Romberg extrapolation, which requires at least two simulations at different time steps Talay & Tubaro (1990)).

2.5.1 Error estimates. The approach we follow is based on the following operator identity (which makes sense on \mathcal{H}^1 for instance, in view of (2.5))

$$\mathcal{L}_\gamma^{-1} = - \int_0^{+\infty} e^{t\mathcal{L}_\gamma} dt.$$

Since

$$\int_{\mathcal{E}} \left(e^{t\mathcal{L}_\gamma} \psi \right) g_\gamma d\mu = \mathbb{E} \left(\psi(q_t, p_t) g_\gamma(q_0, p_0) \right),$$

where the expectation is over all initial conditions (q_0, p_0) distributed according to μ and over all realizations of equilibrium Langevin dynamics (2.1), the leading order correction term (2.29) can be rewritten as

$$\int_{\mathcal{E}} \psi(q, p) f_\gamma(q, p) \mu(dq dp) = - \int_0^{+\infty} \mathbb{E} \left(\psi(q_t, p_t) g_\gamma(q_0, p_0) \right) dt. \quad (2.31)$$

The following result (proved in Section 4.8) shows how to approximate quantities such as (2.31) up to errors $O(\Delta t^\alpha)$, when the invariant measure of the numerical scheme is correct up to terms of order $O(\Delta t^\alpha)$ (as discussed in Section 2.4). The fundamental ingredient is the replacement of the observable ψ by some modified observable (in the spirit of backward analysis). Let us emphasize that we do not require the numerical scheme to be of weak or strong order p in itself. For instance, GLA schemes are only first order correct on trajectories Bou-Rabee & Owahdi (2010), but nonetheless may have invariant measures which are very close to μ . To somewhat simplify the notation, we do not denote explicitly the dependencies in γ (although the reader should keep them in mind).

THEOREM 2.8 Consider a numerical method with an invariant measure $\mu_{\Delta t}$ such that, for $\psi \in \mathcal{S}$,

$$\int_{\mathcal{E}} \psi d\mu_{\Delta t} = \int_{\mathcal{E}} \psi d\mu + \Delta t^\alpha r_{\psi, \Delta t}, \quad (2.32)$$

where the remainder $r_{\psi, \Delta t}$ is uniformly bounded for Δt small enough, and assume that its evolution operator $P_{\Delta t}$ is such that

$$-\frac{\text{Id} - P_{\Delta t}}{\Delta t} = \mathcal{L} + \Delta t S_1 + \cdots + \Delta t^{\alpha-1} S_{\alpha-1} + \Delta t^\alpha \tilde{R}_{\alpha, \Delta t}. \quad (2.33)$$

Then, the integrated correlation of two observables $\psi, \varphi \in \mathcal{S}$ such that

$$\int_{\mathcal{E}} \psi d\mu = \int_{\mathcal{E}} \varphi d\mu = 0, \quad (2.34)$$

can be approximated by a Riemann sum up to an error of order Δt^α :

$$\int_0^{+\infty} \mathbb{E} \left(\psi(q_t, p_t) \varphi(q_0, p_0) \right) dt = \Delta t \sum_{n=0}^{+\infty} \mathbb{E}_{\Delta t} \left(\tilde{\psi}_{\Delta t, \alpha}(q^n, p^n) \varphi(q^0, p^0) \right) + \Delta t^\alpha r_{\Delta t}^{\psi, \varphi}, \quad (2.35)$$

where $r_{\Delta t}^{\psi, \varphi}$ is uniformly bounded for Δt sufficiently small, the expectation $\mathbb{E}_{\Delta t}$ is over all initial conditions (q_0, p_0) distributed according to $\mu_{\Delta t}$, and over all realizations of the Markov chain induced by $P_{\Delta t}$, and the modified observable $\tilde{\psi}_{\Delta t, \alpha}$ is

$$\tilde{\psi}_{\Delta t, \alpha} = \psi_{\Delta t, \alpha} - \int_{\mathcal{E}} \psi_{\Delta t, \alpha} d\mu_{\Delta t}, \quad \psi_{\Delta t, \alpha} = (\text{Id} + \Delta t S_1 \mathcal{L}^{-1} + \cdots + \Delta t^{\alpha-1} S_{\alpha-1} \mathcal{L}^{-1}) \psi. \quad (2.36)$$

In the particular case $\alpha = 2$, which is in fact the most relevant one from a practical viewpoint, it is possible not to modify the observable ψ when the discrete generator is correct at order 2 (see (2.37) below for a precise statement), upon considering a time discretization of the integral which leads to errors of order Δt^2 , for instance a trapezoidal rule. The following result is obtained by an appropriate application of Theorem 2.8 (see Section 4.8 for the proof).

COROLLARY 2.3 (Trapezoidal rule for second order schemes) Consider a numerical scheme whose discrete generator is correct at order 2:

$$-\frac{\text{Id} - P_{\Delta t}}{\Delta t} = \mathcal{L} + \frac{\Delta t}{2} \mathcal{L}^2 + \Delta t^2 \tilde{R}_{\Delta t}. \quad (2.37)$$

Then, for two observables $\varphi, \psi \in \mathcal{S}$ satisfying (2.34),

$$\begin{aligned} & \int_0^{+\infty} \mathbb{E} \left(\psi(q_t, p_t) \varphi(q_0, p_0) \right) dt \\ &= \frac{\Delta t}{2} \mathbb{E}_{\Delta t} \left(\psi_{\Delta t,0}(q^0, p^0) \varphi(q^0, p^0) \right) + \Delta t \sum_{n=1}^{+\infty} \mathbb{E}_{\Delta t} \left(\psi_{\Delta t,0}(q^n, p^n) \varphi(q^0, p^0) \right) + \Delta t^2 r_{\Delta t}^{\psi, \varphi}, \end{aligned} \quad (2.38)$$

where $r_{\Delta t}^{\psi, \varphi}$ is bounded for Δt sufficiently small and

$$\psi_{\Delta t,0} = \psi - \int_{\mathcal{E}} \psi d\mu_{\Delta t}.$$

2.5.2 Numerical approximation. There are two principal ways to estimate the expectations in (2.35) or (2.38), using either several independent realizations of the nonequilibrium dynamics or a single, long trajectory, see for instance the discussion in (Tuckerman, 2010, Section 13.4). When K independent realizations $(q^{n,k}, p^{n,k})$ are generated for N_{iter} time steps each, starting from initial conditions distributed according to $\mu_{\Delta t}$, the expectation in (2.35) may be approximated using empirical averages of the correlation functions as

$$\frac{\Delta t}{K} \sum_{k=1}^K \sum_{n=0}^{N_{\text{iter}}} \left[\psi_{\Delta t, \alpha}(q^{n,k}, p^{n,k}) - \Psi_{\Delta t, \alpha}^{K, N_{\text{iter}}} \right] \varphi(q^{0,k}, p^{0,k}),$$

where $\alpha = 1$ and $\psi_{\Delta t,1} = \psi$ for first order splittings; while $\alpha = 2$ and $\psi_{\Delta t,2} = (1 + \Delta t \mathcal{L}/2)\psi$ for second order ones since $S_1 = \mathcal{L}^2/2$ for the schemes presented in Section 2.2.2 (see for instance (4.25)). The empirical average $\Psi_{\Delta t, \alpha}^{M, N_{\text{iter}}}$ reads

$$\Psi_{\Delta t, \alpha}^{M, N_{\text{iter}}} = \frac{1}{K(1 + N_{\text{iter}})} \sum_{k=1}^K \sum_{n=0}^{N_{\text{iter}}} \psi_{\Delta t, \alpha}(q^{n,k}, p^{n,k}).$$

This formula highlights the other errors arising from the discretization: (i) a statistical error related to the finiteness of K and to the fact that initial conditions are obtained in practice by subsampling a single, long trajectory; (ii) a truncation error related to the finiteness of N_{iter} .

2.5.3 Numerical illustration. We illustrate the convergence results (2.35) and (2.38) for a simple two-dimensional system. We denote $q = (x, y) \in \mathcal{M} = (2\pi\mathbb{T})^2$, and consider the potential

$$V(q) = 2 \cos(2x) + \cos(y).$$

The inverse temperature is fixed to $\beta = 1$ and we consider a trivial mass matrix $M = \text{Id}$ with unit friction $\gamma = 1$. Trajectory data is taken from 10^3 independent runs of fixed time interval 2×10^8 , with the aim to compute the integral of the velocity autocorrelation function. Using the second order $P_{\Delta t}^{\gamma^C, B, A, B, \gamma^C}$ scheme, applying the appropriate correction function (2.38) gives the predicted order Δt^2 result, while the standard Riemann approximation has errors of order Δt . In the numerical results in Figure 1 the corrected approximation gives marginally better results than the trapezoidal rule (though of the same order) due to additional higher order terms being included.

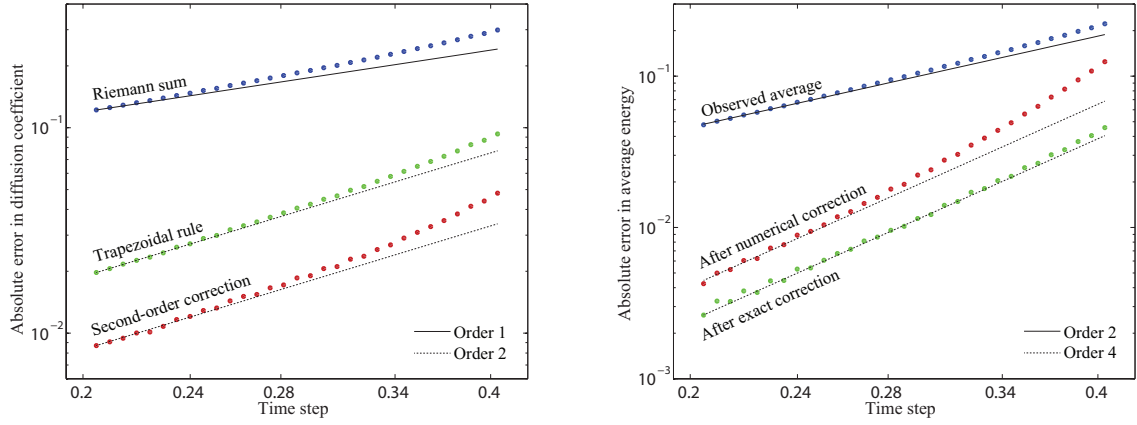


FIG. 1. Left: The error in the value of the integrated velocity autocorrelation function is compared at a number of time steps when computed using a Riemann sum (blue) or the correction term provided in (2.35) (red). The result from computing the integral using the trapezoidal rule (green) is also shown. Right: The error in the computed average of total energy is plotted (blue), with the correction term computed using the same stepsize (red) demonstrating the practical application of the method. We can test the validity of (2.25) in principle by computing the correction more accurately at a smaller time step in a separate simulation, this result is plotted in green. All results are computed using the scheme associated with $P_{\Delta t}^{\gamma C, B, A, B, \gamma C}$ with $\beta = \gamma = 1$.

Let us now numerically confirm the error estimates (2.23)-(2.25)-(2.27). More precisely, we show that, provided the leading correction term (2.29) is estimated by discretizing (2.31) using (2.38) and subtracted from the estimated result, canonical averages are estimated up to errors of order Δt^4 for second order splittings instead of Δt^2 without the correction. We use the same trajectory data as above to approximate the canonical average of the total system energy H . We test the effectiveness of the correction both in practice and principle, by computing the observed average and correction term in the same simulation in the former case, while computing a more accurate correction term independently in the latter case (using a smaller time step $\Delta t = 0.1$). The results are shown in the right panel of Figure 1.

2.6 Overdamped limit

We study in this section the overdamped limit $\gamma \rightarrow +\infty$, assuming that the mass matrix is $M = \text{Id}$. We first study the consistency of the invariant measures of limiting numerical schemes in Section 2.6.1, before stating precise convergence results for second order splitting schemes in Section 2.6.2. We finally relate in Section 2.6.3 the overdamped limit of the correction terms obtained for finite γ to the correction obtained by directly studying the overdamped limit.

2.6.1 Overdamped limits of splitting schemes. The only part of the numerical schemes where the friction enters is the Ornstein-Uhlenbeck process on momenta. The limit $\gamma \rightarrow +\infty$ for $\Delta t > 0$ fixed amounts to resampling momenta according to the Gaussian distribution $\kappa(dp)$ at all time steps. For instance, the numerical scheme associated with the evolution operator $P_{\Delta t}^{\gamma C, B, A, B, \gamma C}$ reduces to

$$q^{n+1} = q^n - \frac{\Delta t^2}{2} \nabla V(q^n) + \frac{\Delta t}{\sqrt{\beta}} G^n, \quad (2.39)$$

where (G^n) are independent and identically distributed standard Gaussian random variables. This is indeed a consistent discretization of the overdamped process (2.8) with an effective time step $h = \Delta t^2/2$, and the invariant measure of this numerical scheme is close to $\bar{\mu}$. Other schemes may have non-trivial large friction limits and invariant measures close to $\bar{\mu}$. This is the case for the scheme associated with the evolution

operator $P_{\Delta t}^{B,A,\gamma C,A,B}$, for which the limiting discrete dynamics reads Leimkuhler & Matthews (2013a)

$$q^{n+1} = q^n - \frac{\Delta t^2}{2} \nabla V(q^n) + \frac{\Delta t}{2\sqrt{\beta}} (G^n + G^{n+1}), \quad (2.40)$$

where (G^n) are independent and identically distributed standard Gaussian random variables. Note that (q^n) is not a Markov chain due to the correlations in the random noises.

On the other hand, the limits of the invariant measures associated with certain schemes are not consistent with the canonical measure $\bar{\mu}$. This is the case for the first-order schemes, as well as the second order splittings listed in item (iii) in Section 2.2.2. For instance, the limit of the scheme associated with $P_{\Delta t}^{\gamma C,A,B}$ reads

$$q^{n+1} = q^n + \frac{\Delta t}{\sqrt{\beta}} G^n.$$

The invariant measure of this Markov chain is the uniform measure on \mathcal{M} , and is therefore very different from the invariant measure $\bar{\mu}$ of the continuous dynamics (2.8) (it amounts to setting $V = 0$). As another example, consider the limit of the scheme associated with $P_{\Delta t}^{\gamma C,B,A}$:

$$q^{n+1} = q^n - \Delta t^2 \nabla V(q^n) + \frac{\Delta t}{\sqrt{\beta}} G^n.$$

This is the Euler-Maruyama discretization of (2.8) with an effective time step $h = \Delta t^2$ but an inverse temperature 2β rather than β .

2.6.2 Rigorous error estimates. The following result quantifies the errors of the invariant measure of second order splitting schemes of Langevin dynamics, for large values of γ . We restrict ourselves to the second order splittings where the Ornstein-Uhlenbeck part is either at the ends or in the middle (categories (i) and (ii) in Section 2.2.2). From a technical viewpoint, we are able here to bound remainder terms uniformly in γ by relying on the properties of the limiting operator $\mathcal{L}_{\text{ovd}}^{-1}$. The result we obtain is the following (see Section 4.9 for the proof).

THEOREM 2.9 Consider any of the second order splittings presented in Section 2.2.2, denote by $\mu_{\gamma,\Delta t}(dq dp)$ its invariant measure, and by $\bar{\mu}_{\gamma,\Delta t}(dq)$ its marginal in the position variable. Then there exists a function $f_{2,\infty} = f_{2,\infty}(q)$ such that, for any smooth function $\psi = \psi(q)$ and $\gamma \geq 1$,

$$\int_{\mathcal{M}} \psi(q) \bar{\mu}_{\gamma,\Delta t}(dq) = \int_{\mathcal{M}} \psi d\bar{\mu} + \Delta t^2 \int_{\mathcal{M}} \psi f_{2,\infty} d\bar{\mu} + r_{\psi,\gamma,\Delta t},$$

where the remainder is of order Δt^4 up to terms exponentially small in $\gamma \Delta t$. More precisely, there exist constants $a, b \geq 0$ and $\kappa > 0$ (depending on ψ) such that

$$|r_{\psi,\gamma,\Delta t}| \leq a \Delta t^4 + b e^{-\kappa \gamma \Delta t}.$$

The expression of $f_{2,\infty}$ depends on the numerical scheme at hand:

$$\begin{aligned} f_{2,\infty}^{\gamma C,B,A,B,\gamma C}(q) &= \frac{1}{8} (-2\Delta V + \beta |\nabla V|^2 + a_{\beta,V}), & a_{\beta,V} &= \int_{\mathcal{M}} \Delta V d\bar{\mu} = \beta \int_{\mathcal{M}} |\nabla V|^2 d\bar{\mu}, \\ f_{2,\infty}^{A,B,\gamma C,B,A}(q) &= -\frac{1}{8} (\Delta V - a_{\beta,V}), \\ f_{2,\infty}^{\gamma C,A,B,A,\gamma C}(q) &= \frac{1}{8} (\Delta V - \beta |\nabla V|^2), \\ f_{2,\infty}^{B,A,\gamma C,A,B}(q) &= 0. \end{aligned} \quad (2.41)$$

Two comments are in order. Note first that the result is stated for observables which depend only on the position variable q since the limiting case $\gamma \rightarrow +\infty$ corresponds to a dynamics on the positions only. There is anyway no restriction in stating the result using such observables since, as already discussed in the

introduction, the error on the marginal in the position variables is the relevant error, momenta being trivial to sample exactly under the canonical measure. Secondly, let us emphasize that the Δt^2 correction term vanishes for the method associated with $P_{\Delta t}^{B,A,\gamma C,A,B}$ (as already noted in Leimkuhler & Matthews (2013a)). This is related to the fact that the corresponding discretization of overdamped Langevin dynamics (formally obtained by setting $\gamma = +\infty$) has an invariant measure which is correct at second-order in the effective time step $h = \Delta t^2/2$.

2.6.3 Overdamped limit of the correction terms. In order to relate the convergence result from Theorem 2.9 to the error estimates from Theorem 2.7, we prove that the limit of the correction functions $f_{2,\gamma}$ as $\gamma \rightarrow +\infty$ agrees with the functions defined in (2.41) (see Section 4.10 for the proof). This can be seen as a statement regarding the permutation of the limits $\gamma \rightarrow +\infty$ and $\Delta t \rightarrow 0$ for the leading correction term, namely, for a function $\psi = \psi(q)$,

$$\begin{aligned} \lim_{\Delta t \rightarrow 0} \lim_{\gamma \rightarrow +\infty} \frac{1}{\Delta t^2} \left(\int_{\mathcal{M}} \psi d\bar{\mu}_{\gamma, \Delta t} - \int_{\mathcal{M}} \psi d\bar{\mu} \right) &= \lim_{\gamma \rightarrow +\infty} \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t^2} \left(\int_{\mathcal{M}} \psi d\bar{\mu}_{\gamma, \Delta t} - \int_{\mathcal{M}} \psi d\bar{\mu} \right) \\ &= \lim_{\gamma \rightarrow +\infty} \int_{\mathcal{M}} \psi (\pi f_{2,\gamma}) d\bar{\mu} \\ &= \int_{\mathcal{M}} \psi f_{2,\infty} d\bar{\mu}. \end{aligned}$$

The result is the following:

PROPOSITION 2.10 There exists a constant $K > 0$ such that, for all $\gamma \geq 1$,

$$\begin{aligned} \left\| f_2^{\gamma C, B, A, B, \gamma C} - \frac{1}{8} (-2\Delta V + \beta |\nabla V|^2 + a_{\beta, V}) \right\|_{H^1(\mu)} &\leq \frac{K}{\gamma}, \\ \left\| f_2^{A, B, \gamma C, B, A} - \frac{1}{8} (-2\Delta V + \beta p^T (\nabla^2 V) p + a_{\beta, V}) \right\|_{H^1(\mu)} &\leq \frac{K}{\gamma}, \\ \left\| f_2^{\gamma C, A, B, A, \gamma C} - \frac{1}{8} (\Delta V - \beta |\nabla V|^2) \right\|_{H^1(\mu)} &\leq \frac{K}{\gamma}, \\ \left\| f_2^{B, A, \gamma C, A, B} - \frac{1}{8} (\Delta V - \beta p^T (\nabla^2 V) p) \right\|_{H^1(\mu)} &\leq \frac{K}{\gamma}, \end{aligned}$$

where the constant $a_{\beta, V}$ is defined in (2.41).

Note that, as expected, the averages with respect to $\kappa(dp)$ of the above limiting functions coincide with the functions $f_{2,\infty}$ given in (2.41), that is, $\pi f_{2,\gamma} = f_{2,\infty} + O(\gamma^{-1})$.

Let us also mention that the overdamped limit of the correction functions $f_{1,\gamma}$ for first order splittings is not well defined. This is not surprising since the invariant measures of the corresponding numerical schemes are not consistent with $\bar{\mu}$, as discussed in Section 2.6.1. For instance, combining (2.10) and the expressions of the correction functions (2.24), we see for instance that there exists a constant $K > 0$ such that

$$\left\| f_1^{\gamma C, B, A} + \frac{\gamma \beta}{2} \mathcal{L}_{\text{ovd}}^{-1} \mathcal{L}_{\text{ovd}, M} V \right\|_{H^1(\mu)} \leq K,$$

where

$$\mathcal{L}_{\text{ovd}, M} = -M^{-1} \nabla V \cdot \nabla_q + \frac{1}{\beta} M : \nabla^2$$

is the generator of the overdamped Langevin dynamics with non-trivial mass matrix:

$$dq_t = -M^{-1} \nabla V(q_t) dt + \sqrt{\frac{2}{\beta}} M^{-1/2} dW_t.$$

Note that, when $M = \text{Id}$, the solution can in fact be analytically computed as $f_1^{\gamma C, B, A} = -\beta(\gamma V + p^T \nabla V)/2$. In any case, $f_1^{\gamma C, B, A}$ diverges as $\gamma \rightarrow +\infty$.

3. Nonequilibrium systems and the computation of transport coefficients

We discuss in this section the numerical estimation of transport properties such as the thermal conductivity, the shear stress, etc. (see Evans & Morriss (2008); Tuckerman (2010) for general physical presentations of the computation of transport coefficients, and (Stoltz, 2012, Section 3.1) for a mathematically oriented introduction).

We consider the prototypical case of the estimation of the autodiffusion coefficient. In this situation, it is relevant to consider nonequilibrium perturbations of the standard equilibrium Langevin dynamics, where some external forcing arising from a constant force $F \in \mathbb{R}^d$ is imposed to the system:

$$\begin{cases} dq_t = M^{-1} p_t dt, \\ dp_t = \left(-\nabla V(q_t) + \eta F \right) dt - \gamma M^{-1} p_t dt + \sqrt{\frac{2\gamma}{\beta}} dW_t. \end{cases} \quad (3.1)$$

We denote by

$$\widetilde{\mathcal{L}} = F \cdot \nabla_p$$

the generator of the perturbation (considered as an operator on $L^2(\mu)$, with domain $H^1(\mu)$). Note that the constant force F does not derive from the gradient of a function on \mathcal{M} . Therefore, the expression of the invariant measure is unknown, but can be obtained as an expansion in powers of η when the magnitude of the forcing is sufficiently small (see Section 3.1). The effect of the force is to create a non-zero average velocity in the direction of F . The magnitude of the average velocity is a property of the system under consideration. For small forcings, it is linear in η , with a constant of proportionality called the *mobility* (see the definition (3.3) below).

We will also be interested in the overdamped limit of the nonequilibrium dynamics (3.1), which reads

$$dq_t = \left(-\nabla V(q_t) + \eta F \right) dt + \sqrt{\frac{2}{\beta}} dW_t. \quad (3.2)$$

The generator of this dynamics is $\mathcal{L}_{\text{ovd}} + \eta \widetilde{\mathcal{L}}_{\text{ovd}}$ with $\widetilde{\mathcal{L}}_{\text{ovd}} = F \cdot \nabla_q$. In this case the physically relevant response turns out to be the average force $-F \cdot \nabla V$ exerted in the direction F .

3.1 Definition of transport coefficients

Following the strategy advertised in Rey-Bellet (2006) (using the kinetic energy as a Lyapunov function), it is easy to show that the dynamics (3.1) has a unique invariant probability measure $\mu_{\gamma, \eta}(dq dp)$ with a smooth density with respect to the Lebesgue measure for any value of $\eta \in \mathbb{R}$. The mobility $v_{F, \gamma}$ is defined as the linear response of the velocity in the direction F as the magnitude of the forcing goes to 0:

$$v_{F, \gamma} = \lim_{\eta \rightarrow 0} \frac{1}{\eta} \int_{\mathcal{E}} F^T M^{-1} p \mu_{\gamma, \eta}(dq dp). \quad (3.3)$$

From linear response theory (see for example the presentation in (Stoltz, 2012, Section 3.1), and the short summary provided in Section 4.11), it can be shown that

$$v_{F, \gamma} = \int_{\mathcal{E}} F^T M^{-1} p f_{0,1, \gamma}(q, p) \mu(dq dp), \quad \mathcal{L}_{\gamma}^* f_{0,1, \gamma} = -\widetilde{\mathcal{L}}^* \mathbf{1} = -\beta F^T M^{-1} p. \quad (3.4)$$

The mobility can therefore be rewritten as the integrated autocorrelation function of the velocity in the direction F :

$$v_{F, \gamma} = \beta \int_0^{+\infty} \mathbb{E} \left[(F^T M^{-1} p_t) (F^T M^{-1} p_0) \right] dt, \quad (3.5)$$

where the expectation is over all initial conditions (q_0, p_0) distributed according to μ and for all realizations of the equilibrium Langevin dynamics (2.1). From this relation, it is easily seen that the mobility is related to the autodiffusion coefficient

$$D_{F, \gamma} = \lim_{t \rightarrow +\infty} \frac{\mathbb{E} \left[(F \cdot (q_t - q_0))^2 \right]}{2t} \quad (3.6)$$

as

$$v_{F,\gamma} = \beta D_{F,\gamma}.$$

In practice, the two most popular ways of estimating a transport coefficient rely on the Green-Kubo formula (3.5) and the linear response of nonequilibrium dynamics in their steady-states (3.3). Since the error estimates for Green-Kubo type formulas have already been discussed in Theorem 2.8, we will restrict ourselves in the sequel to the analysis of the numerical errors introduced by nonequilibrium methods.

3.1.1 Overdamped limit. The overdamped limit of the mobility $v_{F,\gamma}$ is studied in Hairer & Pavliotis (2008), where the authors in fact consider the autodiffusion coefficient $D_{F,\gamma}$. First, it is easily shown that the dynamics (3.2) admits a unique invariant probability measure, which we denote by $\bar{\mu}_\eta(dq)$. The mobility for the overdamped dynamics (3.2) is defined from the linear response of the projected force $-F \cdot \nabla V$ as

$$\bar{v}_F = \lim_{\eta \rightarrow 0} \frac{1}{\eta} \int_{\mathcal{M}} -F^T \nabla V(q) \bar{\mu}_\eta(dq) = \beta \int_{\mathcal{M}} F^T \nabla V(q) \mathcal{L}_{\text{ovd}}^{-1}(F^T \nabla V(q)) \bar{\mu}(dq). \quad (3.7)$$

The derivation of this formula is very similar to the one leading to (3.3). The following result summarizes the limiting behavior of the mobility as the friction increases (recall that we set mass matrices to identity when studying overdamped limits).

LEMMA 3.1 There exists $K > 0$ such that, for any $\gamma \geq 1$,

$$|\gamma v_{F,\gamma} - \bar{v}_F - |F|^2| \leq \frac{K}{\gamma}.$$

This result is already contained in Hairer & Pavliotis (2008), but we nonetheless provide a short alternative proof in Section 4.11.2 (see Remark 4.1 for a more precise comparison of the results). It shows that, in the overdamped regime $\gamma \rightarrow +\infty$,

$$v_{F,\gamma} = \frac{|F|^2 + \bar{v}_F}{\gamma} + O\left(\frac{1}{\gamma^2}\right),$$

which suggests to estimate $v_{F,\gamma}$ using the linear response of $F^T \nabla V$ for large frictions since this quantity is expected to be a good approximation of \bar{v}_F – instead of relying on the standard linear response result (3.3), for which the response is of order $1/\gamma$ and hence difficult to reliably estimate. Error estimates on the numerical approximation are deduced from (3.10) below.

3.2 Numerical schemes for the nonequilibrium Langevin dynamics

We present in this section numerical schemes approximating solutions of (3.1). These schemes reduce to the schemes presented in Section 2.2 when $\eta = 0$. Since the aim is to decompose the evolution generated by $\mathcal{L}_\gamma + \eta \widetilde{\mathcal{L}}$ into analytically integrable parts, there are two principal options: either replace B by

$$B_\eta = B + \eta \widetilde{\mathcal{L}}$$

or replace C by $C + \eta \widetilde{\mathcal{L}}$. However, the schemes built on the latter option do not perform correctly in the overdamped limit since their invariant measures are not consistent with the invariant measures of nonequilibrium overdamped Langevin dynamics (3.2). More precisely, consider for instance the first order scheme generated by $P_{\Delta t}^{A,B,\gamma C + \eta \widetilde{\mathcal{L}}} = e^{\Delta t A} e^{\Delta t B} e^{\Delta t (\gamma C + \eta \widetilde{\mathcal{L}})}$ in the case when $M = \text{Id}$:

$$\begin{cases} q^{n+1} = q^n + \Delta t p^n, \\ \tilde{p}^{n+1} = p^n - \Delta t \nabla V(q^{n+1}), \\ p^{n+1} = \alpha_{\Delta t} \tilde{p}^{n+1} + \frac{1 - \alpha_{\Delta t}}{\gamma} \eta F + \sqrt{\frac{1 - \alpha_{\Delta t}^2}{\beta}} G^n, \end{cases}$$

where $\alpha_{\Delta t}$ is defined after (2.11). As $\gamma \rightarrow +\infty$, a standard Euler-Maruyama discretization of the equilibrium overdamped Langevin dynamics (*i.e.* $\eta = 0$) is obtained, whereas we would like to obtain a consistent discretization of the nonequilibrium overdamped Langevin dynamics (3.2). We therefore rather consider schemes obtained by replacing B with $B + \eta \widetilde{\mathcal{L}}$, such as the first order splitting

$$P_{\Delta t}^{A, B + \eta \widetilde{\mathcal{L}}, \gamma C} = e^{\Delta t A} e^{\Delta t (B + \eta \widetilde{\mathcal{L}})} e^{\gamma \Delta t C},$$

or the second order splitting

$$P_{\Delta t}^{\gamma C, B + \eta \widetilde{\mathcal{L}}, A, B + \eta \widetilde{\mathcal{L}}, C} = e^{\gamma \Delta t C/2} e^{\Delta t (B + \eta \widetilde{\mathcal{L}})/2} e^{\Delta t A} e^{\Delta t (B + \eta \widetilde{\mathcal{L}})/2} e^{\gamma \Delta t C/2}.$$

The numerical scheme associated with the first order splitting scheme $P_{\Delta t}^{A, B + \eta \widetilde{\mathcal{L}}, \gamma C}$

$$\begin{cases} q^{n+1} = q^n + \Delta t p^n, \\ \tilde{p}^{n+1} = p^n - \Delta t \left(\nabla V(q^{n+1}) + \eta F \right), \\ p^{n+1} = \alpha_{\Delta t} \tilde{p}^{n+1} + \sqrt{\frac{1 - \alpha_{\Delta t}^2}{\beta}} G^n, \end{cases}$$

indeed is, in the limit as $\gamma \rightarrow +\infty$, a consistent discretization of the nonequilibrium Langevin dynamics (3.2), and its invariant measure turns out to converge to the invariant measure of (3.2) in the limit $\Delta t \rightarrow 0$.

Following the lines of proof of Proposition 2.3, it can be shown that there exists a unique invariant measure $\mu_{\gamma, \eta, \Delta t}$ for the corresponding Markov chains. The crucial point is that the gradient structure of the force term is never used explicitly in the proofs since we solely rely on the boundedness of the force, so that we are able to obtain convergence results and moment estimates which are independent of the magnitude η of the forcing term provided η is in a bounded subset of \mathbb{R} . We denote below by $P_{\gamma, \eta, \Delta t}$ the evolution operator associated with the numerical schemes.

PROPOSITION 3.1 (Ergodicity of numerical schemes for nonequilibrium systems) Fix $s^* \geq 1$ and $\eta^* > 0$. For any $0 < \gamma < +\infty$, there exists Δt^* such that, for any $0 < \Delta t \leq \Delta t^*$ and $0 \leq \eta \leq \eta^*$, the Markov chain associated with $P_{\gamma, \eta, \Delta t}$ has a unique invariant probability measure $\mu_{\gamma, \eta, \Delta t}$, which admits a density with respect to the Lebesgue measure $dq dp$, and has finite moments: There exists $R > 0$ such that, for any $1 \leq s \leq s^*$,

$$\int_{\mathcal{E}} \mathcal{K}_s d\mu_{\gamma, \eta, \Delta t} \leq R < +\infty,$$

uniformly in the time step Δt and the forcing magnitude η . There also exist $\lambda, K > 0$ (depending on s^*, γ and η^* but not on Δt) such that, for almost all $(q, p) \in \mathcal{E}$, and for all functions $f \in L_{\mathcal{K}_s}^\infty$,

$$\forall n \in \mathbb{N}, \quad \left| \left(P_{\gamma, \eta, \Delta t}^n f \right) (q, p) - \int_{\mathcal{E}} f d\mu_{\gamma, \eta, \Delta t} \right| \leq K \mathcal{K}_s(q, p) e^{-\lambda n \Delta t} \|f\|_{L_{\mathcal{K}_s}^\infty}.$$

Let us emphasize that we do not have any control on the convergence rate λ in terms of η^* , and it could well be that λ goes to 0 as η^* increases.

3.3 Error estimates on transport coefficients from nonequilibrium methods

The following result provides error estimates for the invariant measure for the first order or second order splittings schemes of Section 2.2.2 when B is replaced by B_η .

THEOREM 3.2 Denote by p the order of the splitting scheme, by $f_{\alpha, 0, \gamma}$ the leading order correction function in the case $\eta = 0$ as given by Theorem 2.5 for $\alpha = 1$ and by Theorem 2.7 for $\alpha = 2$. Then, there exists a function $f_{\alpha, 1, \gamma} \in H^1(\mu)$ such that, for any smooth function $\psi \in \mathcal{S}$, there exist $\Delta t^*, \eta^* > 0$ and a constant $K > 0$ for which, for all $0 \leq \eta \leq \eta^*, 0 < \Delta t \leq \Delta t^*$,

$$\int_{\mathcal{E}} \psi d\mu_{\gamma, \eta, \Delta t} = \int_{\mathcal{E}} \psi \left(1 + \eta f_{0, 1, \gamma} + \Delta t^\alpha f_{\alpha, 0, \gamma} + \eta \Delta t^\alpha f_{\alpha, 1, \gamma} \right) d\mu + r_{\psi, \gamma, \eta, \Delta t},$$

where $f_{0,1,\gamma}$ is defined in (3.4), and

$$|r_{\psi,\gamma,\eta,\Delta t}| \leq K(\eta^2 + \Delta t^{\alpha+1}), \quad |r_{\psi,\gamma,\eta,\Delta t} - r_{\psi,\gamma,0,\Delta t}| \leq K\eta(\eta + \Delta t^{\alpha+1}).$$

The proof of this result can be read in Section 4.12. Note that the remainder term now collects higher order terms both as powers of the time-step and the nonequilibrium parameter η . The estimates we obtain on the remainder are however compatible with taking the linear response limit, as made precise by the following error estimate on the transport coefficient (which is an immediate consequence of Theorem 3.2). In order to state the result, we introduce the reference linear response for an observable ψ

$$\mathcal{D}_{\psi,\gamma,0} = \lim_{\eta \rightarrow 0} \frac{1}{\eta} \left(\int_{\mathcal{E}} \psi d\mu_{\gamma,\eta} - \int_{\mathcal{E}} \psi d\mu_{\gamma} \right),$$

and its numerical approximation

$$\mathcal{D}_{\psi,\gamma,\Delta t} = \lim_{\eta \rightarrow 0} \frac{1}{\eta} \left(\int_{\mathcal{E}} \psi d\mu_{\gamma,\eta,\Delta t} - \int_{\mathcal{E}} \psi d\mu_{\gamma,\Delta t} \right).$$

It is often the case that ψ has a vanishing average with respect to μ , as is the case for the function $F^T M^{-1} p$ in (3.3). In general, it however has a non-zero average with respect to the invariant measure $\mu_{\gamma,\Delta t}$ of the numerical scheme associated with a discretization of the equilibrium dynamics.

COROLLARY 3.1 There exist $\Delta t^*, \eta^* > 0$ and a constant $K > 0$ such that, for all $0 \leq \eta \leq \eta^*, 0 < \Delta t \leq \Delta t^*$,

$$\mathcal{D}_{\psi,\gamma,\Delta t} = \mathcal{D}_{\psi,\gamma,0} + \Delta t^\alpha \int_{\mathcal{E}} \psi f_{\alpha,1,\gamma} d\mu + \Delta t^{\alpha+1} r_{\psi,\gamma,\Delta t},$$

where $r_{\psi,\gamma,\Delta t}$ is uniformly bounded.

In particular, we obtain the following estimate on the numerically computed mobility:

$$v_{F,\gamma,\Delta t} = \lim_{\eta \rightarrow 0} \frac{1}{\eta} \left(\int_{\mathcal{E}} F^T M^{-1} p \mu_{\Delta t,\eta}(dq dp) - \int_{\mathcal{E}} F^T M^{-1} p \mu_{\Delta t,0}(dq dp) \right) \quad (3.8)$$

$$= v_{F,\gamma} + \Delta t^\alpha \int_{\mathcal{E}} F^T M^{-1} p f_{\alpha,1,\gamma} d\mu + \Delta t^{\alpha+1} r_{\gamma,\Delta t}, \quad (3.9)$$

where the reference mobility $v_{F,\gamma}$ is defined in (3.4).

3.3.1 Numerical illustration. We consider the same system as in Section 2.5.3, with an external force $F = (1, 0)$ and $K + 1$ forcing strengths $\eta_k = (k - 1)\Delta\eta$ uniformly spaced in the interval $[0, \eta_{\max}]$ with $\eta_{\max} = 0.5$ (so that $\Delta\eta = \eta_{\max}/K$). We fix the friction to $\gamma = 1$ and the inverse temperature to $\beta = 1$. We use a coupling strategy to reduce the statistical noise in the computation of the linear response (3.8). The $K + 1$ replicas of the system are started at the same position $q = (0, 0)$, with the same velocity (sampled according to the canonical measure μ). Each replica experiences the force $-\nabla V + \eta_k F$ (Note that the first replica experiences the reference force $-\nabla V$ corresponding to a discretization of the equilibrium dynamics). Most importantly, the same Gaussian random numbers G^n are used for all replicas to discretize the Brownian motion. Although not carefully documented here, this coupling strategy tremendously decreases the statistical error in the computed linear responses. Such a coupling strategy was already proposed for exclusion processes in Goodman & Lin (2009). However, our experience shows that it fails for higher dimensional systems with more complex potentials (such as Lennard-Jones fluids).

For a given value of the time step Δt , we denote by $(q^{k,n}, p^{k,n})_{n \geq 0}$ the discrete trajectory of the k th replica. The linear response in the projected average velocity δv_{η_k} is approximated over N_{iter} integration steps as

$$\begin{aligned} \delta v_{\eta_k} &= \int_{\mathcal{E}} F^T M^{-1} p \mu_{\Delta t,\eta_k}(dq dp) - \int_{\mathcal{E}} F^T M^{-1} p \mu_{\Delta t,0}(dq dp) \\ &\simeq \frac{1}{N_{\text{iter}}} \sum_{n=1}^{N_{\text{iter}}} F^T M^{-1} (p^{k,n} - p^{1,n}) = \hat{v}_{\eta_k}^{N_{\text{iter}}}. \end{aligned}$$

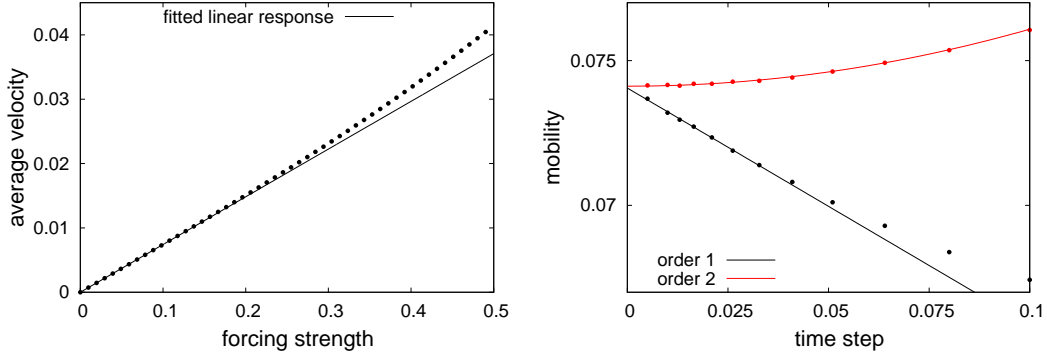


FIG. 2. Left: Linear response of the average velocity δv_η as a function of η ($K = 50$) for the scheme associated with $P_{\Delta t}^{\gamma C, B_\eta, A, B_\eta, \gamma C}$ and $\Delta t = 0.01$, $\gamma = 1$. A linear fit on the first ten values gives $\delta v_\eta \simeq 0.07416\eta$, so that $v_{F, \gamma, \Delta t} = 0.07416$ in this case. Right: Scaling of the mobility $v_{F, \gamma, \Delta t}$ for the first order scheme $P_{\Delta t}^{A, B_\eta, \gamma C}$ and the second order scheme $P_{\Delta t}^{\gamma C, B_\eta, A, B_\eta, \gamma C}$ (with $\gamma = 1$). The fits respectively give $v_{F, \gamma, \Delta t} \simeq 0.0740 + 0.0817\Delta t$ and $v_{F, \gamma, \Delta t} \simeq 0.0741 + 0.197\Delta t^2$.

We then estimate the mobility by a linear fit on the first $K' = 10$ values of $\hat{v}_{\eta_k}^{N_{\text{iter}}}$ considered as a function of η_k (see Figure 2, left). The value $v_{F, \gamma, \Delta t}$ is the estimated slope in the fit. The behavior of the mobility $v_{F, \gamma, \Delta t}$ as a function of the time step is presented in Figure 2 (right) for the numerical schemes associated with the first order splitting $P_{\Delta t}^{A, B_\eta, \gamma C}$ and the second order splitting $P_{\Delta t}^{\gamma C, B_\eta, A, B_\eta, \gamma C}$. We used $N_{\text{iter}} = 4 \times 10^{11}$ for the first order scheme, and $N_{\text{iter}} = 2.5 \times 10^{11}$ for the second order one. The statistical error is very small and error bars are therefore not reported. The computed mobilities can be fitted for small Δt as

$$v_{F, \gamma, \Delta t} \simeq 0.0740 + 0.0817\Delta t$$

for the first-order splitting and

$$v_{F, \gamma, \Delta t} \simeq 0.0741 + 0.197\Delta t^2$$

for the second order splitting scheme, in agreement with the theoretical prediction (3.9).

3.4 Error estimates in the overdamped limit

We now study the numerical errors arising in the simulation of nonequilibrium systems in the large friction limit. We restrict ourselves to the second order splittings where the Ornstein-Uhlenbeck part is either at the ends or in the middle (categories (i) and (ii) in Section 2.2.2). To state the result, we introduce the first order correction to the invariant measure in terms of the magnitude of the nonequilibrium forcing, namely (recall $\mathcal{L}_{\text{ovd}} = F \cdot \nabla_q$)

$$\mathcal{L}_{\text{ovd}}^* f_{0,1,\infty} = -\widetilde{\mathcal{L}}_{\text{ovd}}^* \mathbf{1} = -\beta F^T \nabla V.$$

A simple computation based on (2.10) shows that the functions $f_{0,1,\gamma}$ defined in (3.4) converge in $H^1(\mu)$ to $f_{0,1,\infty}$ (recall that we assume $M = \text{Id}$ in the overdamped regime).

THEOREM 3.3 Denote by $\bar{\mu}_{\gamma, \eta, \Delta t}(dq)$ the marginal of the invariant measure $\mu_{\gamma, \eta, \Delta t}$ of an admissible second order splitting scheme in the position variable, and by $f_{1,0,\infty}$ the leading order correction function in the case $\eta = 0$ as given by Theorem 2.9. Then, there exists a function $f_{2,1,\infty} \in H^1(\mu)$ such that, for any smooth function $\psi \in \mathcal{S}$ depending only on the position variable, there exist $\Delta t^*, \eta^* > 0$ and constants $K, \kappa > 0$ such that, for all $0 \leq \eta \leq \eta^*$, $0 < \Delta t \leq \Delta t^*$ and $\gamma \geq 1$,

$$\int_{\mathcal{M}} \psi(q) \bar{\mu}_{\gamma, \eta, \Delta t}(dq) = \int_{\mathcal{M}} \psi(q) \left(1 + \eta f_{0,1,\infty}(q) + \Delta t^2 f_{2,0,\infty}(q) + \eta \Delta t^2 f_{2,1,\infty}(q) \right) \bar{\mu}(dq) + r_{\psi, \gamma, \eta, \Delta t},$$

with

$$|r_{\psi, \gamma, \eta, \Delta t}| \leq K \left(\eta^2 + \Delta t^3 + e^{-\kappa \gamma \Delta t} \right), \quad |r_{\psi, \gamma, \eta, \Delta t} - r_{\psi, \gamma, 0, \Delta t}| \leq K \eta (\eta + \Delta t^3 + e^{-\kappa \gamma \Delta t}).$$

The proof is presented in Section 4.13. This result allows us to estimate the error in the computation of the transport coefficient $v_{F,\gamma}$ based on (3.7) and Lemma 3.1. Indeed, studying the linear response of the observable $-F^T \nabla V$ and defining

$$\bar{v}_{F,\gamma,\Delta t} = -\lim_{\eta \rightarrow 0} \frac{1}{\eta} \left(\int_{\mathcal{M}} F^T \nabla V(q) \bar{\mu}_{\gamma,\eta,\Delta t}(dq) - \int_{\mathcal{M}} F^T \nabla V(q) \bar{\mu}_{\gamma,\Delta t}(dq) \right),$$

it holds

$$\bar{v}_F = \bar{v}_{F,\gamma,\Delta t} + \Delta t^2 \int_{\mathcal{M}} F^T \nabla V(q) f_{2,1,\infty}(q) \bar{\mu}(dq) + r_{\Psi,\gamma,\Delta t},$$

with $|r_{\Psi,\gamma,\Delta t}| \leq a(\Delta t^3 + e^{-\kappa\gamma\Delta t})$ for some $a > 0$. Therefore,

$$v_{F,\gamma} = \frac{|F|^2 + \bar{v}_F}{\gamma} + O\left(\frac{1}{\gamma^2}\right) = \frac{|F|^2 + \bar{v}_{F,\gamma,\Delta t}}{\gamma} + O\left(\frac{1}{\gamma^2}, \frac{\Delta t^2}{\gamma}, \frac{e^{-\kappa\gamma\Delta t}}{\gamma}\right). \quad (3.10)$$

In the latter expression, $\bar{v}_{F,\gamma,\Delta t}$ can be numerically estimated, as documented at the end of Section 3.3.

4. Proof of the results

Unless otherwise stated, the default norm $\|f\|$ and scalar product $\langle f, g \rangle$ are the ones associated with the Hilbert space $L^2(\mu)$. Adjoint operators are also by default considered as adjoints on $L^2(\mu)$. Recall that

$$C = -\frac{1}{\beta} \nabla_p^* \nabla_p = -\frac{1}{\beta} \sum_{i=1}^N \sum_{\alpha=1}^d \partial_{p_{i,\alpha}}^* \partial_{p_{i,\alpha}}, \quad (4.1)$$

with $p_i = (p_{i,1}, \dots, p_{i,d})$ since $\partial_{p_{i,\alpha}}^* = -\partial_{p_{i,\alpha}} + \beta p_{i,\alpha}$.

4.1 Large friction behavior of \mathcal{L}_γ^{-1}

The proof of Lemma 2.1 follows the same lines as the proof of uniform hypocoercive estimates in the corrected version of (Joubaud & Stoltz, 2012a, Theorem 3) (see the erratum Joubaud & Stoltz (2013) or the updated preprint version Joubaud & Stoltz (2012b)). We however reproduce a simplified version of it for completeness.

Proof of Lemma 2.1. We show that the operator \mathcal{L}_γ is uniformly hypocoercive for $\gamma \geq 1$, provided the domain of the operator is restricted to \mathcal{H}_\perp^1 . To this end, we decompose \mathcal{L}_γ for $\gamma \geq 1$ as

$$\mathcal{L}_\gamma = \mathcal{L}_1 + (\gamma - 1)C.$$

The proof of (Hairer & Pavliotis, 2008, Theorem 6.2) shows that there exists $\tilde{\alpha} > 0$ such that, for all smooth functions $u \in \mathcal{H}_\perp^1$,

$$-\langle u, \mathcal{L}_1 u \rangle \geq \tilde{\alpha} \langle u, u \rangle,$$

where the norm induced by $\langle \cdot, \cdot \rangle$ is equivalent to the $H^1(\mu)$ norm. More precisely, $\langle \cdot, \cdot \rangle$ is the bilinear form defined by

$$\langle u, v \rangle = a \langle u, v \rangle + b \langle \nabla_p u, \nabla_p v \rangle - \langle \nabla_p u, \nabla_q v \rangle - \langle \nabla_q u, \nabla_p v \rangle + b \langle \nabla_q u, \nabla_q v \rangle,$$

with appropriate coefficients $a \gg b \gg 1$. It follows that there exists $\alpha > 0$ independent of γ such that

$$\alpha \|u\|_{H^1(\mu)}^2 - (\gamma - 1) \langle u, Cu \rangle \leq -\langle u, \mathcal{L}_\gamma u \rangle. \quad (4.2)$$

Let us now show that

$$\forall u \in \mathcal{H}_\perp^1, \quad -\langle u, Cu \rangle \geq 0. \quad (4.3)$$

Using the rewriting (4.1) of the operator C , and the commutation relations $[\partial_{p_{i,\alpha}}, \partial_{p_{j,\alpha'}}^*] = \beta \delta_{\alpha,\alpha'} \delta_{ij}$, a simple computation shows

$$\begin{aligned} \langle \langle u, (\partial_{p_{i,\alpha}})^* \partial_{p_{i,\alpha}} u \rangle \rangle &= (a + \beta b) \|\partial_{p_{i,\alpha}} u\|^2 + b \|\nabla_p \partial_{p_{i,\alpha}} u\|^2 \\ &\quad + b \|\nabla_q \partial_{p_{i,\alpha}} u\|^2 - 2 \langle \nabla_q \partial_{p_{i,\alpha}} u, \nabla_p \partial_{p_{i,\alpha}} u \rangle - \beta \langle \partial_{q_{i,\alpha}} u, \partial_{p_{i,\alpha}} u \rangle \\ &\geq \left(a + \beta \left(b - \frac{1}{2} \right) \right) \|\partial_{p_{i,\alpha}} u\|^2 + (b - 1) \|\nabla_p \partial_{p_{i,\alpha}} u\|^2 \\ &\quad + (b - 1) \|\nabla_q \partial_{p_{i,\alpha}} u\|^2 - \frac{\beta}{2} \|\partial_{q_{i,\alpha}} u\|^2. \end{aligned} \quad (4.4)$$

Now, since the Gaussian measure $\kappa(dp)$ satisfies a Poincaré inequality, there exists a constant $A > 0$ such that, for all $i = 1, \dots, N$ and $\alpha = 1, \dots, d$,

$$\|\partial_{q_{i,\alpha}} u\|^2 \leq A \|\nabla_p \partial_{q_{i,\alpha}} u\|^2. \quad (4.5)$$

Note indeed that $\partial_{q_{i,\alpha}} u$ has a vanishing average with respect to the Gaussian measure $\kappa(dq)$ because

$$\int_{\mathbb{R}^{dN}} \partial_{q_{i,\alpha}} u(q, p) \kappa(dp) = \partial_{q_{i,\alpha}} \bar{u}(q) = 0$$

for functions $u \in \mathcal{H}_\perp^1$. Therefore,

$$\sum_{i=1}^N \sum_{\alpha=1}^d \|\partial_{q_{i,\alpha}} u\|^2 \leq A \sum_{i,j=1}^N \sum_{\alpha,\alpha'=1}^d \|\partial_{p_{j,\alpha'}} \partial_{q_{i,\alpha}} u\|^2 = A \sum_{j=1}^N \sum_{\alpha'=1}^d \|\nabla_q \partial_{p_{j,\alpha'}} u\|^2.$$

Summing (4.4) on $i \in \{1, \dots, N\}$ and $\alpha \in \{1, \dots, d\}$, the quantity (4.3) is seen to be non-negative for an appropriate choice of constants $a \gg b \gg 1$.

From (4.2), we then deduce that there exists a constant $K > 0$ such that, for any $\gamma \geq 1$ and for any $u \in \mathcal{H}_\perp^1 \cap H^2(\mu)$, it holds $\|u\|_{H^1(\mu)} \leq K \|\mathcal{L}_\gamma u\|_{H^1(\mu)}$. Taking inverses and passing to the limit in \mathcal{H}_\perp^1 gives

$$\forall \gamma \geq 1, \quad \forall u \in \mathcal{H}_\perp^1, \quad \|\mathcal{L}_\gamma^{-1} u\|_{H^1(\mu)} \leq K \|u\|_{H^1(\mu)},$$

which is the desired result. \square

We are now in position to write the

Proof of Theorem 2.2. We write the proof for \mathcal{L}_γ^{-1} . The estimates for $(\mathcal{L}_\gamma^*)^{-1}$ are obtained by using the self-adjointness up to a unitary transform $\mathcal{L}_\gamma^* = \mathcal{R} \mathcal{L}_\gamma \mathcal{R}$ (the momentum reversal operator being defined in (2.4)), and the fact that $\mathcal{R} C \mathcal{R} = C$, $\mathcal{R} \mathcal{L}_{\text{ovd}} \mathcal{R} = \mathcal{L}_{\text{ovd}}$ and $\mathcal{R}(A + B) \mathcal{R} = -(A + B)$.

The lower bound in (2.9) could be obtained directly provided V is not constant, by considering the special case

$$\mathcal{L}_\gamma \left(p^T \nabla V + \gamma(V - v) \right) = p^T M^{-1} (\nabla^2 V) p - |\nabla V|^2,$$

where v is a constant chosen such that $p^T \nabla V + \gamma(V - v)$ has a vanishing average with respect to μ . This example is also useful to motivate the fact that, in general, solutions of the Poisson equation $\mathcal{L}_\gamma u_\gamma = f$ have divergent parts of order γ as $\gamma \rightarrow +\infty$.

Let us now turn to the refined upper and lower bounds (2.10), which we prove using techniques from asymptotic analysis. Consider $f \in \mathcal{H}^1$, and $u_\gamma \in \mathcal{H}^1$ the unique solution of the following Poisson equation $\mathcal{L}_\gamma u_\gamma = f$. The above example suggests the following expansion in inverse powers of γ :

$$u_\gamma = \gamma u^{-1} + u^0 + \frac{1}{\gamma} u^1 + \dots \quad (4.6)$$

To rigorously prove this expansion, we first proceed formally, taking (4.6) as an ansatz, plugging it into $\mathcal{L}_\gamma u = f$ and identifying terms according to powers of γ . This leads to

$$\begin{aligned} C u^{-1} &= 0, \\ (A + B) u^{-1} + C u^0 &= 0, \\ (A + B) u^0 + C u^1 &= f. \end{aligned}$$

The first equality implies that $u^{-1} = u^{-1}(q)$ since C satisfies a Poincaré inequality on $L^2(\kappa)$ (where κ is defined in (1.5)). The second then reduces to $Cu^0 = -M^{-1}p \cdot \nabla_q u^{-1}$, from which we deduce $u^0(q, p) = p^T \nabla u^{-1}(q) + \tilde{u}^0(q)$. Inserting this expression in the third equality gives

$$Cu^1 = f - p^T M^{-1} (\nabla^2 u^{-1}) p - p^T M^{-1} \nabla \tilde{u}^0 + (\nabla V)^T \nabla u^{-1}.$$

The solvability condition for this equation is that the right-hand side has a vanishing average with respect to κ , i.e. belongs to the kernel of π . This condition reads

$$\frac{1}{\beta} \Delta u^{-1} - (\nabla V)^T \nabla u^{-1} = \pi f,$$

so that $u^{-1} = \mathcal{L}_{\text{ovd}}^{-1} \pi f$ (which is well defined since πf has a vanishing average with respect to $\bar{\mu}$). Note that the function u^{-1} is in $H^{n+2}(\bar{\mu})$ when $f \in H^n(\mu)$ (by elliptic regularity, using also the fact that $e^{-\beta V(q)}$ is a smooth function bounded from above and below on \mathcal{M}), so that $p^T M^{-1} (\nabla^2 u^{-1}) p$ belongs to $L^2(\mu)$. The equation determining u^1 then reduces to

$$Cu^1 = (f - \pi f) - p^T M^{-1} \nabla \tilde{u}^0 - p^T M^{-1} (\nabla^2 u^{-1}) p + \frac{1}{\beta} \Delta u^{-1}.$$

Since $C(p^T A p) = -p^T M^{-1} (A + A^T) p + 2\beta^{-1} \text{Tr}(A)$, we can choose

$$u^1(q, p) = [C^{-1}(f - \pi f)](q, p) + \frac{1}{2} p^T (\nabla^2 u^{-1}(q)) p + p^T \nabla_q \tilde{u}^0(q).$$

Coming back to (4.6), we see that the proposed approximate solution is such that

$$\mathcal{L}_\gamma \left(u_\gamma - \gamma u^{-1} - u^0 - \frac{1}{\gamma} u^1 \right) = -\frac{1}{\gamma} (A + B) u^1. \quad (4.7)$$

We now choose \tilde{u}^0 such that $(A + B)u^1$ belongs to \mathcal{H}_\perp^1 , which amounts to

$$\pi(A + B)p^T \nabla_q \tilde{u}^0 = \mathcal{L}_{\text{ovd}} \tilde{u}^0 = -\pi(A + B)C^{-1}(f - \pi f).$$

It is easily checked that $\tilde{u}^0 = -\mathcal{L}_{\text{ovd}}^{-1} \pi(A + B)C^{-1}(f - \pi f)$ is a well defined element in $H^1(\mu)$ for $f \in H^1(\mu)$.

Combining (4.7) and Lemma 2.1, we see that there exists a constant $R > 0$, such that, for all $\gamma \geq 1$, it holds $\|u_\gamma - \gamma u^{-1} - u^0\|_{H^1(\mu)} \leq R \|f\|_{H^1(\mu)} / \gamma$ for the above choices of functions u^{-1}, u^0 . This indeed gives (2.10). \square

4.2 Ergodicity results for numerical schemes

Proof of Lemma 2.2. We write the proof for the scheme associated with the evolution operator $P_{\Delta t}^{B,A,\gamma C}$, starting by the case $s = 1$, before turning to the general case $s \geq 2$. The proofs for other schemes are very similar, and we therefore skip them.

The numerical scheme corresponding to $P_{\Delta t}^{B,A,\gamma C}$ is (2.11). We introduce $m \in (0, +\infty)$ such that $m \leq M \leq m^{-1}$ (in the sense of symmetric matrices). A simple computation shows that

$$\begin{aligned} \mathbb{E} \left[(p^{n+1})^2 \middle| \mathcal{F}_n \right] &= (p^n - \Delta t \nabla V(q^n))^T \alpha_{\Delta t}^2 (p^n - \Delta t \nabla V(q^n)) + \frac{1}{\beta} \text{Tr} \left[(1 - \alpha_{\Delta t}^2) M^2 \right] \\ &\leq e^{-2m\gamma\Delta t} (p^n)^2 + 2\Delta t \|\nabla V\|_{L^\infty} |p^n| + \Delta t^2 \|\nabla V\|_{L^\infty}^2 + \frac{1 - e^{-2\gamma\Delta t/m}}{\beta m^2} \\ &\leq \left(e^{-2m\gamma\Delta t} + \varepsilon \Delta t \right) (p^n)^2 + \Delta t \left(\frac{1}{\varepsilon} + \Delta t \right) \|\nabla V\|_{L^\infty}^2 + \frac{1 - e^{-2\gamma\Delta t/m}}{\beta m^2}. \end{aligned}$$

We choose for instance $\varepsilon = m\gamma$, in which case

$$0 \leq a_{\Delta t} = e^{-2m\gamma\Delta t} + \varepsilon \Delta t \leq \exp \left(-\frac{m\gamma\Delta t}{2} \right)$$

and

$$0 \leq \tilde{b}_{\Delta t} = \Delta t \left(\frac{1}{\varepsilon} + \Delta t \right) \|\nabla V\|_{L^\infty}^2 + \frac{1 - e^{-2\gamma\Delta t/m}}{\beta m^2} \leq \Delta t \left[\frac{2}{m\gamma} \|\nabla V\|_{L^\infty}^2 + \frac{4\gamma}{\beta m^3} \right]$$

for Δt sufficiently small. Finally, since $\mathcal{K}_2(q, p) = 1 + |p|^2$,

$$\mathbb{E} [\mathcal{K}_2(q^{n+1}, p^{n+1}) \mid \mathcal{F}_n] \leq a_{\Delta t} \mathcal{K}_2(q^n, p^n) + 1 - a_{\Delta t} + \tilde{b}_{\Delta t}.$$

We therefore define $b_{\Delta t} = 1 - a_{\Delta t} + \tilde{b}_{\Delta t}$. It is easily checked that (2.15) holds.

To obtain (2.16), we iterate the bound (2.15):

$$\begin{aligned} P_{\Delta t}^n \mathcal{K}_s &\leq a_{\Delta t}^n \mathcal{K}_s + b_{\Delta t} (1 + a_{\Delta t} + \dots + a_{\Delta t}^{n-1}) = a_{\Delta t}^n \mathcal{K}_s + \frac{b_{\Delta t}}{1 - a_{\Delta t}} \\ &\leq \exp(-C_a n \Delta t) \mathcal{K}_s + \frac{C_b \Delta t}{1 - \exp(-C_a \Delta t)} \leq \exp(-C_a n \Delta t) \mathcal{K}_s + \frac{2C_b}{C_a}, \end{aligned}$$

provided Δt is sufficiently small, which is the desired inequality.

The computations are similar for a general power $s \geq 2$. We write $p^{n+1} = \alpha_{\Delta t} p^n + \delta_{\Delta t}$ with $\delta_{\Delta t} = -\alpha_{\Delta t} \Delta t \nabla V(q^n) + \sqrt{\beta^{-1}(1 - \alpha_{\Delta t}^2) M G^n}$. Note that $\delta_{\Delta t}$ is of order $\Delta t^{1/2}$ because of the random term. We work componentwise, using the assumption that M is diagonal, so that, denoting by m_i the mass of the i th degree of freedom,

$$\begin{aligned} (p_i^{n+1})^{2s} &= \left(e^{-\gamma\Delta t/m_i} p_i^n + \delta_{i,\Delta t} \right)^{2s} \\ &= e^{-2s\gamma\Delta t/m_i} (p_i^n)^{2s} + 2s e^{-(2s-1)\gamma\Delta t/m_i} (p_i^n)^{2s-1} \delta_{i,\Delta t} \\ &\quad + s(2s-1) e^{-2(s-1)\gamma\Delta t/m_i} (p_i^n)^{2(s-1)} \delta_{i,\Delta t}^2 + \dots \end{aligned}$$

Taking expectations,

$$\begin{aligned} \mathbb{E} \left[(p_i^{n+1})^{2s} \mid \mathcal{F}_n \right] &= e^{-2s\gamma\Delta t/m_i} (p_i^n)^{2s} - 2s \Delta t e^{-2s\gamma\Delta t/m_i} (p_i^n)^{2s-1} \partial_{q_i} V(q^n) \\ &\quad + s(2s-1) e^{-2(s-1)\gamma\Delta t/m_i} (p_i^n)^{2(s-1)} \left(\Delta t^2 e^{-2\gamma\Delta t/m_i} \partial_{q_i}^2 V(q^n) + \frac{(1 - e^{-2\gamma\Delta t/m_i}) m_i}{\beta} \right) \\ &\quad + \Delta t^2 r_{s,\Delta t,i}(q^n) \left(1 + (p_i^n)^{2s-3} \right), \end{aligned}$$

where the remainder $r_{s,\Delta t,i}(q^n)$ is uniformly bounded as $\Delta t \rightarrow 0$. Distinguishing between $|p_i| \geq 1/\varepsilon$ and $|p_i| \leq 1/\varepsilon$, we have

$$|p_i|^{2s-m} \leq \varepsilon^m (p_i)^{2s} + \frac{1}{\varepsilon^{2s-m}},$$

from which we obtain

$$\mathbb{E} \left[(p_i^{n+1})^{2s} \mid \mathcal{F}_n \right] \leq \hat{a}_{\Delta t, \varepsilon, i} (p_i^n)^{2s} + \hat{b}_{\Delta t, \varepsilon, i},$$

with

$$\begin{aligned} \hat{a}_{\Delta t, \varepsilon, i} &= e^{-2s\gamma\Delta t/m_i} + 2s\varepsilon\Delta t \|\partial_{q_i} V\|_{L^\infty} \\ &\quad + s(2s-1)\varepsilon^2 \left(\Delta t^2 \|\partial_{q_i}^2 V\|_{L^\infty} + \frac{(1 - e^{-2\gamma\Delta t/m_i}) m_i}{\beta} \right) + \varepsilon^3 \Delta t^2 \|r_{s,\Delta t,i}\|_{L^\infty}, \end{aligned}$$

and

$$\begin{aligned} \hat{b}_{\Delta t, \varepsilon, i} &= \frac{2s}{\varepsilon} \Delta t \|\partial_{q_i} V\|_{L^\infty} \\ &\quad + \frac{s(2s-1)}{\varepsilon^2} \left(\Delta t^2 \|\partial_{q_i}^2 V\|_{L^\infty} + \frac{(1 - e^{-2\gamma\Delta t/m_i}) m_i}{\beta} \right) + \Delta t^2 \left(1 + \frac{1}{\varepsilon^3} \right) \|r_{s,\Delta t,i}\|_{L^\infty}. \end{aligned}$$

The proof is then concluded as in the case $s = 1$ by choosing ε sufficiently small (independently of Δt). \square

Proof of Lemma 2.3. It is sufficient to prove the result for indicator functions of Borel sets $A = A_q \times A_p \subset \mathcal{E}$ (where $A_q \subset \mathcal{M}$ while $A_p \subset \mathbb{R}^{dN}$). We therefore aim at proving

$$\mathbb{P}((q^n, p^n) \in A \mid |p^0| \leq p_{\max}) \geq \alpha \nu(A)$$

for a well chosen probability measure ν and a constant $\alpha > 0$. The idea of the proof is to explicitly rewrite q^n and p^n as perturbations of the reference evolution corresponding to $\nabla V = 0$ and $(q^0, p^0) = (0, 0)$. Since we consider smooth potentials and the position space is compact, the perturbation can be uniformly controlled when the initial momenta are within a compact set.

We write the proof for the scheme associated with the evolution operator $P_{\Delta t}^{B, A, \gamma C}$, as in the proof of Lemma 2.2. A simple computation shows that, for $n \geq 1$,

$$q^n = q^0 + \Delta t M^{-1} (p^{n-1} + \dots + p^0) - \Delta t^2 M^{-1} (\nabla V(q^{n-1}) + \dots + \nabla V(q^0))$$

and

$$\begin{aligned} p^n &= \alpha_{\Delta t}^n p^0 - \Delta t \alpha_{\Delta t} (\nabla V(q^{n-1}) + \alpha_{\Delta t} \nabla V(q^{n-2}) + \dots + \alpha_{\Delta t}^{n-1} \nabla V(q^0)) \\ &\quad + \sqrt{\frac{1 - \alpha_{\Delta t}^2}{\beta}} M (G^{n-1} + \alpha_{\Delta t} G^{n-2} + \dots + \alpha_{\Delta t}^{n-1} G^0). \end{aligned}$$

Denote by \mathcal{G}^n the centered Gaussian random variable

$$\mathcal{G}^n = \sqrt{\frac{1 - \alpha_{\Delta t}^2}{\beta}} M (G^{n-1} + \alpha_{\Delta t} G^{n-2} + \dots + \alpha_{\Delta t}^{n-1} G^0).$$

Introduce also

$$\begin{aligned} F^n &= -\alpha_{\Delta t} (\nabla V(q^{n-1}) + \alpha_{\Delta t} \nabla V(q^{n-2}) + \dots + \alpha_{\Delta t}^{n-1} \nabla V(q^0)), \\ \mathcal{P}^n &= \alpha_{\Delta t}^n p^0 + \Delta t F^n, \\ \mathcal{Q}^n &= q^0 + \Delta t M^{-1} \left(\Delta t \sum_{m=0}^{n-1} F^m + \frac{1 - \alpha_{\Delta t}^n}{1 - \alpha_{\Delta t}} p^0 \right) - \Delta t^2 M^{-1} (\nabla V(q^{n-1}) + \dots + \nabla V(q^0)). \end{aligned}$$

With this notation,

$$p^n = \mathcal{P}^n + \mathcal{G}^n, \quad q^n = \mathcal{Q}^n + \tilde{\mathcal{G}}^n,$$

where

$$\begin{aligned} \tilde{\mathcal{G}}^n &= \Delta t M^{-1} \sum_{m=1}^{n-1} \mathcal{G}^m \\ &= \Delta t \sqrt{\frac{1 - \alpha_{\Delta t}^2}{\beta}} M^{-1} (G^{n-2} + (1 + \alpha_{\Delta t}) G^{n-3} + \dots + (1 + \alpha_{\Delta t} + \dots + \alpha_{\Delta t}^{n-2}) G^0) \end{aligned}$$

is a centered Gaussian random variable. Now,

$$\mathbb{P}((q^n, p^n) \in A \mid |p^0| \leq p_{\max}) = \mathbb{P}((\tilde{\mathcal{G}}^n, \mathcal{G}^n) \in (A_q - \mathcal{Q}^n) \times (A_p - \mathcal{P}^n) \mid |p^0| \leq p_{\max}). \quad (4.8)$$

In fact, we consider in the sequel that the random variable $\tilde{\mathcal{G}}^n$ has values in \mathbb{R}^{dN} rather than \mathcal{M} and understand $A_q - \mathcal{Q}^n$ as a subset of \mathbb{R}^{dN} rather than \mathcal{M} . This amounts to neglecting the possible periodic images, and henceforth reduces the probability on the right-hand side of the above inequality. This is however not a problem since we seek a lower bound.

Note that $\Delta t F^n$ is uniformly bounded: using $0 \leq \alpha_{\Delta t} \leq \exp(-\gamma m \Delta t)$ in the sense of symmetric, positive matrices (with $m \leq M \leq m^{-1}$),

$$|\Delta t F^n| \leq \sqrt{dN} \|\nabla V\|_{L^\infty} \frac{\Delta t}{1 - \exp(-\gamma m \Delta t)} \leq \frac{2\sqrt{dN}}{m\gamma} \|\nabla V\|_{L^\infty}$$

provided Δt is sufficiently small. Therefore, there exists a constant $R > 0$ (depending on p_{\max}) and $\Delta t^* > 0$ such that, for all time steps $0 < \Delta t \leq \Delta t^*$ and corresponding integration steps $0 \leq n \leq T/\Delta t$,

$$|\mathcal{Q}^n| \leq R, \quad |\mathcal{P}^n| \leq R. \quad (4.9)$$

A lengthy but straightforward computation shows that the variance of the centered Gaussian vector $(\tilde{\mathcal{G}}^n, \mathcal{G}^n)$ is

$$\mathcal{V}^n = \mathbb{E} \left[\left(\tilde{\mathcal{G}}^n, \mathcal{G}^n \right)^T \left(\tilde{\mathcal{G}}^n, \mathcal{G}^n \right) \right] = \begin{pmatrix} \mathcal{V}_{qq}^n & \mathcal{V}_{qp}^n \\ \mathcal{V}_{qp}^n & \mathcal{V}_{pp}^n \end{pmatrix}$$

with

$$\begin{cases} \mathcal{V}_{qq}^n = \frac{\Delta t (1 - \alpha_{\Delta t}^2)}{(1 - \alpha_{\Delta t})^2} M^{-1} \left((n-1)\Delta t - \frac{2\Delta t \alpha_{\Delta t}}{1 - \alpha_{\Delta t}} (1 - \alpha_{\Delta t}^{n-1}) + \frac{\Delta t \alpha_{\Delta t}^2}{1 - \alpha_{\Delta t}^2} (1 - \alpha_{\Delta t}^{2(n-1)}) \right), \\ \mathcal{V}_{qp}^n = \frac{\Delta t \alpha_{\Delta t}}{\beta (1 - \alpha_{\Delta t})} \left(1 - \alpha_{\Delta t}^{n-1} (1 + \alpha_{\Delta t}) + \alpha_{\Delta t}^{2n-1} \right), \\ \mathcal{V}_{pp}^n = \frac{M}{\beta} (1 - \alpha_{\Delta t}^{2n}). \end{cases}$$

To check that this expression is appropriate, we note that it converges as $\Delta t \rightarrow 0$ with $n\Delta t \rightarrow T$ to the variance of the limiting continuous process

$$dq_t = M^{-1} p_t dt, \quad dp_t = -\gamma M^{-1} p_t dt + \sqrt{\frac{2\gamma}{\beta}} dW_t,$$

starting from $(q_0, p_0) = (0, 0)$, which reads

$$\mathcal{V} = \begin{pmatrix} \mathcal{V}_{qq} & \mathcal{V}_{qp} \\ \mathcal{V}_{qp} & \mathcal{V}_{pp} \end{pmatrix},$$

with

$$\begin{cases} \mathcal{V}_{qq} = \frac{1}{\beta \gamma} \left(2T - \frac{M}{\gamma} (3 - 4\alpha_T + \alpha_T^2) \right), \\ \mathcal{V}_{qp} = \frac{M}{\beta \gamma} (1 - \alpha_T)^2, \\ \mathcal{V}_{pp} = \frac{M}{\beta} (1 - \alpha_T^2). \end{cases}$$

Upon reducing $\Delta t^* > 0$, it holds $\mathcal{V}/2 \leq \mathcal{V}^{\lceil T/\Delta t \rceil} \leq 2\mathcal{V}$ for $0 < \Delta t \leq \Delta t^*$. In particular, $\mathcal{V}^{\lceil T/\Delta t \rceil}$ is invertible for T sufficiently large. For a set $E_q \times E_p \subset \mathbb{R}^{2dN}$, it then holds

$$\begin{aligned} \mathbb{P} \left(\left(\tilde{\mathcal{G}}^{\lceil T/\Delta t \rceil}, \mathcal{G}^{\lceil T/\Delta t \rceil} \right) \in E \right) &= (2\pi)^{-dN} \det \left(\mathcal{V}^{\lceil T/\Delta t \rceil} \right)^{-1/2} \int_{E_q \times E_p} \exp \left(-\frac{1}{2} x^T \left(\mathcal{V}^{\lceil T/\Delta t \rceil} \right)^{-1} x \right) dx \\ &\geq \pi^{-dN} 2^{-3dN/2} \det(\mathcal{V})^{-1/2} \int_{E_q \times E_p} \exp(-x^T \mathcal{V}^{-1} x) dx. \end{aligned} \quad (4.10)$$

The results follows by combining (4.8)-(4.9)-(4.10) and introducing the probability measure

$$\nu(A_q \times A_p) = Z_R^{-1} \inf_{|\mathcal{Q}|, |\mathcal{P}| \leq R} \int_{(A_q - \mathcal{Q}) \times (A_p - \mathcal{P})} \exp(-x^T \mathcal{V}^{-1} x) dx,$$

as well as $\alpha = Z_R \pi^{-dN} 2^{-3dN/2} \det(\mathcal{V})^{-1/2}$. \square

Proof of Proposition 2.3. We only prove (2.18) and (2.17) since the other results are standard. To obtain the bound (2.18), we first note that, by the results of Hairer & Mattingly (2011), there exists $\tilde{\lambda} > 0$ such

that, for any function $f \in L^\infty_{\mathcal{K}_s, \Delta t}$ and $0 < \Delta t \leq \Delta t^*$ (the critical time step being given by Lemmas 2.2 and 2.3),

$$\forall m \in \mathbb{N}, \quad \left| \left(\left[P_{\Delta t}^{[T/\Delta t]} \right]^m f \right) (q, p) \right| \leq K \mathcal{K}_s(q, p) e^{-\tilde{\lambda} m} \|f\|_{L^\infty_{\mathcal{K}_s}}.$$

For a general index $n \in \mathbb{N}$, we write

$$n = m_n \left\lceil \frac{T}{\Delta t} \right\rceil + \tilde{n}, \quad 0 \leq \tilde{n} \leq \left\lceil \frac{T}{\Delta t} \right\rceil - 1,$$

so that, using the contractivity property $|P_{\Delta t} f(q, p)| \leq |f(q, p)|$,

$$|P_{\Delta t}^n f(q, p)| \leq K \mathcal{K}_s(q, p) e^{-\tilde{\lambda} m_n} \|f\|_{L^\infty_{\mathcal{K}_s}}.$$

Introducing $\lambda = \tilde{\lambda}/T$, the argument of the exponent reads

$$\tilde{\lambda} m_n = \lambda (n - \tilde{n}) \Delta t \frac{T}{\Delta t} \left\lceil \frac{T}{\Delta t} \right\rceil^{-1} \geq \frac{\lambda n \Delta t}{2} - \lambda T,$$

when Δt is sufficiently small. This gives (2.18).

The moment estimate (2.17) is obtained by averaging (2.15) with respect to the invariant measure:

$$\int_{\mathcal{E}} (P_{\Delta t} \mathcal{K}_s) d\mu_{\gamma, \Delta t} \leq a_{\Delta t} \int_{\mathcal{E}} \mathcal{K}_s d\mu_{\gamma, \Delta t} + b_{\Delta t}.$$

Since $\mu_{\gamma, \Delta t}$ is invariant,

$$\int_{\mathcal{E}} (P_{\Delta t} \mathcal{K}_s) d\mu_{\gamma, \Delta t} = \int_{\mathcal{E}} \mathcal{K}_s d\mu_{\gamma, \Delta t},$$

so that

$$(1 - a_{\Delta t}) \int_{\mathcal{E}} \mathcal{K}_s d\mu_{\gamma, \Delta t} \leq b_{\Delta t},$$

which is the desired result. \square

4.3 Some useful results

4.3.1 Functional estimates on \mathcal{L}_γ^{-1} . The aim of this section is to give some estimates on the norm of the operator \mathcal{L}_γ^{-1} in the weighted Sobolev spaces $\tilde{H}^m(\mu)$ and $\tilde{W}_{\mathcal{K}_s}^{m, \infty}$. The notation $\tilde{\mathcal{F}}$ for a functional space \mathcal{F} refers to the subspace composed of functions with vanishing average with respect to μ :

$$\tilde{\mathcal{F}} = \left\{ \psi \in \mathcal{F} \mid \int_{\mathcal{E}} \psi d\mu = 0 \right\}.$$

The proof is a slight adaption of the techniques from Talay (2002). This adaption is motivated by the fact one of the assumptions made in (Talay, 2002, Hypothesis 1.1) (which treats a more general dynamics than the one we consider) is that $\nabla_p H$ should be bounded. This is not the case here since $\nabla_p H = M^{-1}p$. In addition, we want to highlight more precisely the operator norms which are implicit in Talay (2002).

Let us first emphasize that many of the arguments in Talay (2002) can be drastically simplified in our setting since the position space \mathcal{M} is compact. The first point we want to mention is that the commutator technique used in Talay (2002) is based on convergence estimates in $L^2(\mu)$, as given by (Talay, 2002, Lemma 3.2). Such estimates are a consequence of hypocoercivity estimates in $\tilde{H}^1(\mu)$. We can then proceed with the remainder of the proofs until the end of (Talay, 2002, Section 3.3) since the only technical assumption required to this end is a good control of the Hessian $\nabla^2 H$. We then obtain the boundedness of \mathcal{L}_γ^{-1} as an operator on any weighted Sobolev space $\tilde{H}^m(\mu)$. Pointwise estimates on compact subsets of \mathcal{E} follow thanks to Sobolev embeddings and the fact that $e^{-\beta H}$ is uniformly bounded from above and below on compact sets. More precisely, there exist $m_* \geq 1$ and $\alpha_* > 0$ such that, for any compact subset $\mathcal{C} \subset \mathcal{E}$, there exists a constant $R_{\mathcal{C}} > 0$ such that

$$\forall (q, p) \in \mathcal{C}, \quad \left| e^{t \mathcal{L}_\gamma} \psi(q, p) \right| \leq R_{\mathcal{C}} e^{-\alpha_* t} \|\psi\|_{H^{m_*}(\mu)}. \quad (4.11)$$

Similar estimates hold for the derivatives of $e^{t\mathcal{L}_\gamma}\psi$ (upon increasing the value of m_*).

Some modifications are however required to obtain bounds over the full configuration space \mathcal{E} , namely bounds in the spaces $\tilde{W}_{\mathcal{K}_s}^{m_*,\infty}$. Indeed, the boundedness of $\nabla_p H$ is used in the proof of (Talay, 2002, Lemma 3.12). We can however prove a very similar statement.

LEMMA 4.1 Introduce the weight functions

$$\Pi_s(q, p) = \left(\frac{1}{1 + |p|^2} \right)^s.$$

There exists an integer $s_* \geq 1$ (sufficiently large), such that, for any $s \geq s_*$, there are real numbers $R_s, \alpha_s > 0$ for which

$$\int_{\mathcal{E}} |e^{t\mathcal{L}_\gamma}\psi|^2 \Pi_s \leq R_s e^{-\alpha_s t} \left(\int_{\mathcal{E}} |\psi|^2 \Pi_s + \|\psi\|_{H^{m_*}(\mu)} \right).$$

The estimate of Lemma 4.1 can be extended to account for derivatives of $e^{t\mathcal{L}_\gamma}\psi$ of arbitrary order (as sketched in Talay (2002)). This finally gives the following L^∞ estimate thanks to a Sobolev embedding and the control of $\|\psi\|_{H^{m_*}(\mu)}$ by $\|\psi\|_{W_{\mathcal{K}_r}^{m_*,\infty}}$ (for any $r \geq 1$).

PROPOSITION 4.1 For any integer $m \geq 1$, there exist $s_m \geq 1$ and $p \geq 1$ such that, for any $s \geq s_m$, there is $R > 0$ for which

$$\forall \psi \in \tilde{W}_{\mathcal{K}_s}^{m,\infty}, \quad \|\mathcal{L}_\gamma^{-1}\psi\|_{W_{\mathcal{K}_s}^{m,\infty}} \leq R \|\psi\|_{W_{\mathcal{K}_s}^{p,\infty}}. \quad (4.12)$$

Let us conclude this section with the

Proof of Lemma 4.1. The proof is based on a Gronwall inequality: Denoting by $u(t) = e^{t\mathcal{L}_\gamma}\psi$,

$$\frac{d}{dt} \left(\frac{1}{2} \int_{\mathcal{E}} |u(t)|^2 \Pi_s \right) = \int_{\mathcal{E}} (\mathcal{L}_\gamma u(t)) u(t) \Pi_s = \int_{\mathcal{E}} u(t) \mathcal{L}_\gamma^\dagger (u(t) \Pi_s),$$

where

$$\mathcal{L}_\gamma^\dagger = -(A + B) + \gamma \operatorname{div}_p \left(M^{-1} p \cdot + \frac{1}{\beta} \nabla_p \cdot \right)$$

is the adjoint of \mathcal{L}_γ on the flat space $L^2(dq dp)$. Note that

$$\mathcal{L}_\gamma^\dagger(fg) = f \mathcal{L}_\gamma^\dagger g + g \mathcal{L}_\gamma^\dagger f + \frac{2\gamma}{\beta} \nabla_p f \cdot \nabla_p g - \gamma \operatorname{Tr}(M^{-1}) fg.$$

Therefore,

$$\begin{aligned} \frac{d}{dt} \left(\frac{1}{2} \int_{\mathcal{E}} |u(t)|^2 \Pi_s \right) &= \int_{\mathcal{E}} |u(t)|^2 \mathcal{L}_\gamma^\dagger \Pi_s + \int_{\mathcal{E}} u(t) (\mathcal{L}_\gamma^\dagger u(t)) \Pi_s \\ &\quad + \frac{2\gamma}{\beta} \int_{\mathcal{E}} u(t) \nabla_p u(t) \cdot \nabla_p \Pi_s - \gamma \operatorname{Tr}(M^{-1}) \int_{\mathcal{E}} |u(t)|^2 \Pi_s. \end{aligned}$$

An integration by parts shows that, for any function $W(p)$ with values in \mathbb{R}^{dN} ,

$$\int_{\mathcal{E}} u(t) \nabla_p u(t) \cdot W = - \int_{\mathcal{E}} u(t) \operatorname{div}_p (u(t) W) = - \int_{\mathcal{E}} u(t) \nabla_p u(t) \cdot W - \int_{\mathcal{E}} |u(t)|^2 \operatorname{div}_p(W),$$

so that

$$\int_{\mathcal{E}} u(t) \nabla_p u(t) \cdot W = - \frac{1}{2} \int_{\mathcal{E}} |u(t)|^2 \operatorname{div}_p(W).$$

A similar treatment gives

$$\int_{\mathcal{E}} u(t) ((A + B)u(t)) \Pi_s = - \frac{1}{2} \int_{\mathcal{E}} |u(t)|^2 (A + B) \Pi_s = \frac{1}{2} \int_{\mathcal{E}} |u(t)|^2 \nabla V(q) \cdot \nabla_p \Pi_s.$$

We therefore have

$$\begin{aligned} \frac{d}{dt} \left(\frac{1}{2} \int_{\mathcal{E}} |u(t)|^2 \Pi_s \right) &= \int_{\mathcal{E}} |u(t)|^2 \mathcal{L}_\gamma^\dagger \Pi_s - \frac{1}{2} \int_{\mathcal{E}} |u(t)|^2 \nabla V(q) \cdot \nabla_p \Pi_s - \frac{\gamma}{2} \int_{\mathcal{E}} |u(t)|^2 \operatorname{div}_p (M^{-1} p \Pi_s) \\ &\quad - \frac{\gamma}{2\beta} \int_{\mathcal{E}} |u(t)|^2 \Delta_p \Pi_s - \frac{\gamma}{\beta} \int_{\mathcal{E}} |\nabla_p u(t)|^2 \Pi_s. \end{aligned} \quad (4.13)$$

Since

$$\partial_{p_i} \Pi_s = -2s \frac{p_i}{1+|p|^2} \Pi_s,$$

it is easily checked that

$$\begin{aligned} \mathcal{L}_\gamma^\dagger \Pi_s &= \nabla V \cdot \nabla_p \Pi_s + \gamma \left(\operatorname{Tr}(M^{-1}) - 2s \frac{p^T M^{-1} p}{1+|p|^2} \right) \Pi_s + \frac{\gamma}{\beta} \Delta_p \Pi_s \\ &= \frac{1}{2} \nabla V \cdot \nabla_p \Pi_s + \gamma \left(\operatorname{Tr}(M^{-1}) - 2s \frac{p^T M^{-1} p}{1+|p|^2} \right) \Pi_s + h_s \Pi_s, \end{aligned}$$

where h_s is a bounded function going to 0 as $|p| \rightarrow +\infty$. Moreover,

$$\operatorname{div}_p (M^{-1} p \Pi_s) = \left(\operatorname{Tr}(M^{-1}) - 2s \frac{p^T M^{-1} p}{1+|p|^2} \right) \Pi_s.$$

Therefore, forgetting about the non-negative last term in (4.13),

$$\frac{d}{dt} \left(\frac{1}{2} \int_{\mathcal{E}} |u(t)|^2 \Pi_s \right) \leq \frac{\gamma}{2} \int_{\mathcal{E}} |u(t)|^2 \left(\operatorname{Tr}(M^{-1}) - 2s \frac{p^T M^{-1} p}{1+|p|^2} \right) \Pi_s + \int_{\mathcal{E}} |u(t)|^2 \tilde{h}_s \Pi_s,$$

where \tilde{h}_s still is bounded a function going to 0 as $|p| \rightarrow +\infty$. There exists $s_* \geq 1$ (large enough) and $P_* > 0$ such that

$$\frac{\gamma}{2} \left(\operatorname{Tr}(M^{-1}) - 2s_* \frac{p^T M^{-1} p}{1+|p|^2} \right) + \tilde{h}_s \leq -\frac{\gamma s_*}{2m_{\max}}$$

for $|p| \geq P_*$ (where $M \leq m_{\max} \operatorname{Id}$ in the sense of symmetric matrices). For these choices,

$$\begin{aligned} \frac{d}{dt} \left(\frac{1}{2} \int_{\mathcal{E}} |u(t)|^2 \Pi_s \right) &\leq -\frac{\gamma s_*}{2m_{\max}} \int_{\mathcal{E}} |u(t)|^2 \Pi_s \\ &\quad + \int_{|p| \leq P_*} |u(t)|^2 \left[\frac{\gamma}{2} \left(\operatorname{Tr}(M^{-1}) - 2s_* \frac{p^T M^{-1} p}{1+|p|^2} \right) + \frac{\gamma s_*}{2m_{\max}} + \tilde{h}_s \right] \Pi_s. \end{aligned}$$

We use the estimate (4.11) to obtain an exponential decay of the second integral on the right-hand side, and conclude by a Gronwall estimate. \square

4.3.2 Expansion of the evolution operator. We give in this section an expression of the evolution operator

$$P_t = e^{tA_M} \dots e^{tA_1}$$

which can easily be compared to the evolution operator $e^{t(A_1 + \dots + A_M)}$. We assume that all the elementary evolution semigroups e^{tA_i} are well defined on some reference Banach space. This is the case for instance when the operators A_i satisfy the conditions of the Hille-Yosida theorem Pazy (1983). All the operator equalities stated in this section have to be considered in the strong sense, namely $T_1 = T_2$ means $T_1 \varphi = T_2 \varphi$ for all $\varphi \in \mathcal{S} \subset D(T_1) \cap D(T_2)$. It is easy to check that the operators A, B, C acting on Banach spaces $W_{\mathcal{H}_s}^{m, \infty}$ or Hilbert spaces $H^m(\mu)$ satisfy the conditions of the Hille-Yosida theorem (upon adding a positive constant to them). It is in fact possible to analytically write down the action of the associated semigroups:

$$\begin{cases} (e^{tA} \varphi)(q, p) = \varphi(q + tM^{-1}p, p), \\ (e^{tB} \varphi)(q, p) = \varphi(q, p - t\nabla V(q)), \\ (e^{tC} \varphi)(q, p) = \int_{\mathbb{R}^{dN}} \varphi \left(e^{-\gamma M^{-1}t} p + \left(\frac{1 - e^{-2\gamma M^{-1}t}}{\beta} M \right)^{1/2} x \right) \frac{e^{-|x|^2/2}}{(2\pi)^{dN/2}} dx. \end{cases} \quad (4.14)$$

The key building block for the subsequent numerical analysis is the following equality:

$$P_t = P_0 + t \left. \frac{dP_t}{dt} \right|_{t=0} + \frac{t^2}{2} \left. \frac{d^2 P_t}{dt^2} \right|_{t=0} + \cdots + \frac{t^n}{n!} \left. \frac{d^n P_t}{dt^n} \right|_{t=0} + \frac{t^{n+1}}{n!} \int_0^1 (1-\theta)^n \left. \frac{d^{n+1} P_s}{ds^{n+1}} \right|_{s=\theta t} d\theta.$$

Now,

$$\begin{aligned} \frac{dP_t}{dt} &= A_M e^{tA_M} \cdots e^{tA_1} + e^{tA_M} A_{M-1} e^{tA_{M-1}} \cdots e^{tA_1} + \cdots + e^{tA_M} \cdots e^{tA_1} A_1 \\ &= \mathcal{T}[(A_1 + \cdots + A_M)P_t] \end{aligned}$$

where \mathcal{T} indicates the ordering operator which exchanges operators so that the operators with the smallest indices (or their associated semigroups) are farthest to the right. In fact, simple computations show that

$$\frac{d^n P_t}{dt^n} = \mathcal{T}[(A_1 + \cdots + A_M)^n P_t].$$

Therefore, the following equality holds when applied to sufficiently smooth functions:

$$\begin{aligned} P_t &= \text{Id} + t(A_1 + \cdots + A_M) + \frac{t^2}{2} \mathcal{T}[(A_1 + \cdots + A_M)^2] + \cdots + \frac{t^n}{n!} \mathcal{T}[(A_1 + \cdots + A_M)^n] \\ &\quad + \frac{t^{n+1}}{n!} \int_0^1 (1-\theta)^n \mathcal{T}[(A_1 + \cdots + A_M)^{n+1} P_{\theta t}] d\theta. \end{aligned} \quad (4.15)$$

4.3.3 BCH formula. It is important to rewrite the various terms in the right-hand side of (4.15) under a form more amenable to analytical computations. More precisely, it is particularly convenient to write

$$\mathcal{T}[(A_1 + \cdots + A_M)^n] = (A_1 + \cdots + A_M)^n + S_n,$$

where the operator S_n involves commutators $[A_i, A_j]$. In fact, the expressions of the operators S_n can be obtained from the BCH formula for first order splittings (see for instance (Hairer *et al.*, 2006, Section III.4.2)): for $M = 3$,

$$e^{\Delta t A_3} e^{\Delta t A_2} e^{\Delta t A_1} = e^{\Delta t \mathcal{A}}, \quad \mathcal{A} = A_1 + A_2 + A_3 + \frac{\Delta t}{2} ([A_3, A_1 + A_2] + [A_2, A_1]) + \cdots,$$

and from the symmetric BCH formula for second order involving 3 operators (obtained by composition of the standard BCH formula involving 2 operators):

$$e^{\Delta t A_1/2} e^{\Delta t A_2/2} e^{\Delta t A_3} e^{\Delta t A_2/2} e^{\Delta t A_1/2} = e^{\Delta t \mathcal{A}}, \quad (4.16)$$

with

$$\begin{aligned} \mathcal{A} &= A_1 + A_2 + A_3 + \frac{\Delta t^2}{12} \left([A_3, [A_3, A_2]] + [A_2 + A_3, [A_2 + A_3, A_1]] \right. \\ &\quad \left. - \frac{1}{2} [A_2, [A_2, A_3]] - \frac{1}{2} [A_1, [A_1, A_2 + A_3]] \right) + \cdots \end{aligned} \quad (4.17)$$

where we do not write down the expressions of the higher order terms Δt^{2n} (for $n \geq 2$). Let us insist that these formulas are only formal (since the operators appearing the argument of the exponential on the right-hand side involve more and more derivatives), but nonetheless allow us to find the algebraic expressions of S_n upon formally expanding the exponential as

$$e^{\Delta t \mathcal{A}} = \text{Id} + \Delta t \mathcal{A} + \frac{\Delta t^2}{2} \mathcal{A}^2 + \cdots$$

and identifying terms with the same powers of Δt in (4.15).

4.3.4 Approximate inverse operators. Consider an invertible operator A , and a perturbation $\Delta t^\alpha B$ for some exponent $\alpha \geq 1$. In the typical situations encountered in this article, $\Delta t^\alpha B$ may be thought of to be small since it is of order Δt^α , although B may not be bounded with respect to A since it may involve higher order derivatives than A does. It is therefore impossible in general to properly define the inverse of $A + \Delta t^\alpha B$.

However, it is possible to define an approximate inverse, which we define as an operator $Q_{\Delta t, n}$ such that

$$(A + \Delta t^\alpha B)Q_{\Delta t, n} = \text{Id} + \Delta t^{(n+1)\alpha} R_{\Delta t, n}.$$

To this end we truncate the formal series expression of the inverse of the operator $A + \Delta t^\alpha B = A(\text{Id} + \Delta t^\alpha A^{-1}B)$ in powers of $A^{-1}B$:

$$A^{-1} - \Delta t^\alpha A^{-1}BA^{-1} + \Delta t^{2\alpha} A^{-1}BA^{-1}BA^{-1} + \dots$$

For instance, $Q_{\Delta t, 1} = A^{-1} - \Delta t^\alpha A^{-1}BA^{-1}$ and $Q_{\Delta t, 2} = A^{-1} - \Delta t^\alpha A^{-1}BA^{-1} + \Delta t^{2\alpha} A^{-1}BA^{-1}BA^{-1}$.

4.4 Proof of Theorem 2.5

We write the proof for the scheme associated with $P_{\Delta t}^{\gamma C, B, A} = e^{\gamma \Delta t C} e^{\Delta t B} e^{\Delta t A}$, the proof for the scheme $P_{\Delta t}^{\gamma C, A, B}$ following the same lines. The results for the other schemes are then obtained with the TU lemma (Lemma 2.4). Without loss of generality, we perform the proof for a function ψ with average 0 with respect to μ (recovering the general case by adding a constant to ψ in the final expression).

PROOF OF (2.23). Using the results of Section 4.3, a simple computation shows that

$$P_{\Delta t}^{\gamma C, B, A} = \text{Id} + \Delta t \mathcal{L}_\gamma + \frac{\Delta t^2}{2} (\mathcal{L}_\gamma^2 + S_1) + \Delta t^3 R_{1, \Delta t}, \quad S_1 = [C, A + B] + [B, A], \quad (4.18)$$

where the subscript index 1 refers to the order of the splitting. More precisely,

$$R_{1, \Delta t} = \frac{1}{2} \int_0^1 (1 - \theta)^2 \mathcal{R}_{\theta \Delta t} d\theta,$$

where \mathcal{R}_s is a finite linear combination of terms of the form $C^\gamma e^{sC} B^\beta e^{sB} A^\alpha e^{sA}$ with $\alpha, \beta, \gamma \geq 0$ and $\alpha + \beta + \gamma = 3$. In any case, $R_{1, \Delta t}$ involves at most 6 derivatives. Therefore,

$$\begin{aligned} & \int_{\mathcal{E}} \left[\left(\frac{\text{Id} - P_{\Delta t}^{\gamma C, B, A}}{\Delta t} \right) \varphi \right] (1 + \Delta t f_{1, \gamma}) d\mu \\ &= - \int_{\mathcal{E}} \left[\left(\mathcal{L}_\gamma + \frac{\Delta t}{2} (\mathcal{L}_\gamma^2 + S_1) + \Delta t^2 R_{1, \Delta t} \right) \varphi \right] (1 + \Delta t f_{1, \gamma}) d\mu \\ &= - \Delta t \int_{\mathcal{E}} \left(\frac{1}{2} S_1 \varphi + (\mathcal{L}_\gamma \varphi) f_{1, \gamma} \right) d\mu - \Delta t^2 \int_{\mathcal{E}} \left(\left[\frac{1}{2} (\mathcal{L}_\gamma^2 + S_1) \varphi \right] f_{1, \gamma} + (R_{1, \Delta t} \varphi) (1 + \Delta t f_{1, \gamma}) \right) d\mu. \end{aligned}$$

The dominant term on the right-hand side can be written as

$$\int_{\mathcal{E}} \left(\frac{1}{2} S_1 \varphi + (\mathcal{L}_\gamma \varphi) f_{1, \gamma} \right) d\mu = \int_{\mathcal{E}} \varphi \left[\frac{1}{2} S_1^* \mathbf{1} + \mathcal{L}_\gamma^* f_{1, \gamma} \right] d\mu,$$

which suggests to choose

$$\mathcal{L}_\gamma^* f_{1, \gamma} = -\frac{1}{2} S_1^* \mathbf{1}. \quad (4.19)$$

The function $f_{1, \gamma}$ is well defined since the right-hand side belongs to \mathcal{H}^1 . A direct computation indeed shows that $S_1^* \mathbf{1} \in H^1(\mu)$ (see (4.23)). The centering condition follows from the fact that $\mathbf{1} \in \text{Ker}(S_1)$: using the adjoint operator,

$$\int_{\mathcal{E}} S_1^* \mathbf{1} d\mu = \int_{\mathcal{E}} S_1 \mathbf{1} d\mu = 0.$$

With the choice (4.19),

$$\begin{aligned} \int_{\mathcal{E}} \left[\left(\frac{\text{Id} - P_{\Delta t}^{\mathcal{C},B,A}}{\Delta t} \right) \varphi \right] (1 + \Delta t f_{1,\gamma}) d\mu \\ = -\Delta t^2 \int_{\mathcal{E}} \left(\left[\frac{1}{2} (\mathcal{L}_{\gamma}^2 + S_1) \varphi \right] f_{1,\gamma} + (R_{1,\Delta t} \varphi) (1 + \Delta t f_{1,\gamma}) \right) d\mu. \end{aligned} \quad (4.20)$$

We would like, at this stage, to replace the observable φ appearing on the left hand side by the function

$$\left(\frac{\text{Id} - P_{\Delta t}^{\mathcal{C},B,A}}{\Delta t} \right)^{-1} \psi.$$

However, we do not have any control on the derivatives of this function (Corollary 2.1 allows to control the norm of the function, not of its derivatives), whereas such a control is required to bound the remainder terms. We therefore consider the approximate inverse operator (see Section 4.3.4)

$$Q_{1,\Delta t} = -\mathcal{L}_{\gamma}^{-1} + \frac{\Delta t}{2} (\text{Id} + \mathcal{L}_{\gamma}^{-1} S_1 \mathcal{L}_{\gamma}^{-1}),$$

which is such that

$$\left(\frac{\text{Id} - P_{\Delta t}^{\mathcal{C},B,A}}{\Delta t} \right) Q_{1,\Delta t} = - \left(\mathcal{L}_{\gamma} + \frac{\Delta t}{2} (\mathcal{L}_{\gamma}^2 + S_1) + \Delta t^2 R_{1,\Delta t} \right) Q_{1,\Delta t} = \text{Id} + \Delta t^2 Z_{1,\Delta t},$$

and replace φ by $Q_{1,\Delta t} \psi$ in (4.20) (Note that $\varphi = Q_{1,\Delta t} \psi$ is a well defined function since ψ has average 0 with respect to μ). This gives

$$\int_{\mathcal{E}} \psi (1 + \Delta t f_{1,\gamma}) d\mu = \Delta t^2 \int_{\mathcal{E}} \left[(\tilde{R}_{1,\Delta t} \psi) f_{1,\gamma} + \hat{R}_{1,\Delta t} \psi \right] d\mu \quad (4.21)$$

where the functions $\tilde{R}_{1,\Delta t} \psi, \hat{R}_{1,\Delta t} \psi$ belong to \mathcal{S} when ψ does.

On the other hand, by definition of the invariant measure $\mu_{\gamma,\Delta t}$,

$$\int_{\mathcal{E}} \left(\frac{\text{Id} - P_{\Delta t}^{\mathcal{C},B,A}}{\Delta t} \right) \varphi d\mu_{\gamma,\Delta t} = 0,$$

so that, replacing again φ by $Q_{1,\Delta t} \psi$,

$$\int_{\mathcal{E}} \psi d\mu_{\gamma,\Delta t} = -\Delta t^2 \int_{\mathcal{E}} (Z_{1,\Delta t} \psi) d\mu_{\gamma,\Delta t}. \quad (4.22)$$

The right-hand side is indeed well defined and uniformly bounded for small Δt since $Z_{1,\Delta t} \psi \in \mathcal{S}$ and the moments of $\mu_{\gamma,\Delta t}$ are uniformly bounded for Δt sufficiently small (see Proposition 2.3).

The combination of (4.21) and (4.22) gives (2.23) for the splitting scheme $P_{\Delta t}^{\mathcal{C},B,A}$.

PROOF OF (2.24). The function $f_1^{\mathcal{C},B,A}$ (denoted by $f_{1,\gamma}$ above) is determined by the equation

$$\mathcal{L}_{\gamma}^* f_1^{\mathcal{C},B,A} = -\frac{1}{2} S_1^* \mathbf{1} = -\frac{1}{2} ([C, A+B] + [B, A])^* \mathbf{1},$$

where we have used $[\mathcal{L}_{\gamma}^2]^* \mathbf{1} = 0$ to simplify the right-hand side. Now, $[C, A+B]^* = [C, A+B]$ since $C^* = C$ and $(A+B)^* = -(A+B)$. Therefore, $[C, A+B]^* \mathbf{1} = 0$. In addition,

$$[B, A]^* \mathbf{1} = -(A+B)^* g = (A+B)g$$

since $A^* = -A + g$ and $B^* = -B - g$. Therefore,

$$S_1^* \mathbf{1} = (A+B)g. \quad (4.23)$$

This gives the first expression in (2.24).

To obtain the expressions of $f_1^{A,\gamma^C,B}$ and f_1^{B,A,γ^C} , we use the TU lemma, where the operators $U_{\Delta t}$ respectively read $e^{\gamma\Delta t C}e^{\Delta t B} = \text{Id} + \Delta t(B + \gamma C) + \Delta t^2 R_{\Delta t}$ and $e^{\gamma\Delta t C}$ (which preserves μ). We actually are in a situation similar to (2.22):

$$f_1^{B,A,\gamma^C} = f_1^{\gamma^C,B,A}, \quad f_1^{A,\gamma^C,B} = f_1^{\gamma^C,B,A} + B^* \mathbf{1}.$$

The expressions for the first order corrections when the operators A and B are exchanged are obtained by noting that the sign of $S_1^* \mathbf{1}$ is changed and that $f_1^{B,\gamma^C,A} = f_1^{\gamma^C,A,B} + A^* \mathbf{1}$.

4.5 Proof of Proposition 2.6

We use a very standard strategy: first, we propose an ansatz for the correction term $f_{1,\gamma}$ as

$$f_{1,\gamma} = f_1^0 + \gamma f_1^1 + \gamma^2 f_1^2 + \dots,$$

then identify the two leading order terms in this expression, and finally use the resolvent estimate of Theorem 2.1 to conclude. Note that our ansatz is not obvious since the estimate of Theorem 2.1 shows that, in general, a leading order correction term of order $1/\gamma$ should be considered. It turns out however that, due to the specific structure of the right-hand side of (2.24) (namely the fact that the right-hand is at leading order in γ the image by the Hamiltonian operator of some function), such a divergent leading order term is not necessary.

Consider for instance the case when $f_{1,\gamma}$ is $f_1^{\gamma^C,B,A}$. This function solves

$$\left[-(A+B) + \gamma C \right] f_1^{\gamma^C,B,A} = -\frac{1}{2}(A+B)g,$$

so that we consider the ansatz $f_1^{\gamma^C,B,A} = g/2 + \gamma f_1^1 + \dots$. Identifying terms with same powers of γ , we see that the correction term f_1^1 should satisfy

$$(A+B)f_1^1 = \frac{1}{2}Cg = \frac{\beta}{2}p^T M^{-2} \nabla V.$$

Possible solutions are defined up to elements of the kernel of $A+B$ (which contains function of the form $\varphi \circ H$). One possible choice is to set $f_1^1 = \beta p^T M^{-2} p/4$, in which case

$$\mathcal{L}_\gamma^* \left(f_1^{\gamma^C,B,A} - \frac{g}{2} - \gamma f_1^1 \right) = \gamma^2 C f_1^1.$$

In view of (2.1), this implies that there exists a constant $K > 0$ such that

$$\left\| f_1^{\gamma^C,B,A} - \frac{g}{2} - \gamma f_1^1 \right\|_{L^2(\mu)} \leq K\gamma$$

for $\gamma \leq 1$, which gives the desired estimate on $f_1^{\gamma^C,B,A}$. Similar computations give the estimate on $f_1^{\gamma^C,A,B}$, while the estimates on the remaining functions are obtained from (2.24).

4.6 Proof of Theorem 2.7

The proof follows the same lines as the proof for the first order splitting schemes (see Section 4.4). We present only the required modifications. We write the proof for $P_{\Delta t}^{\gamma^C,B,A,B,\gamma^C}$ since the correction term has a much simpler right-hand side than $P_{\Delta t}^{A,B,\gamma^C,B,A}$.

PROOF OF (2.25). Expanding up to terms of order Δt^5 the formal expression of $P_{\Delta t}^{\gamma^C,B,A,B,\gamma^C}$ given by the BCH expansion (4.16), we obtain the following equality

$$P_{\Delta t}^{\gamma^C,B,A,B,\gamma^C} = \text{Id} + \Delta t(\mathcal{L}_\gamma + \Delta t^2 S_2) + \frac{\Delta t^2}{2} \left(\mathcal{L}_\gamma^2 + \Delta t^2 (\mathcal{L}_\gamma S_2 + S_2 \mathcal{L}_\gamma) \right) + \frac{\Delta t^3}{6} \mathcal{L}_\gamma^3 + \frac{\Delta t^4}{24} \mathcal{L}_\gamma^4 + \Delta t^5 R_{2,\Delta t},$$

where

$$R_{2,\Delta t} = \frac{1}{24} \int_0^1 (1-\theta)^4 \mathcal{R}_{\theta\Delta t} d\theta,$$

\mathcal{R}_s being a finite linear combination of terms of the form $C^\gamma e^{sC} B^\beta e^{sB} A^\alpha e^{sA}$ with $\alpha, \beta, \gamma \geq 0$ and $\alpha + \beta + \gamma = 5$; and

$$S_2 = \frac{1}{12} (S_{2,0} + \gamma S_{2,1} + \gamma^2 S_{2,2}), \quad (4.24)$$

with

$$\begin{cases} S_{2,0} = [A, [A, B]] - \frac{1}{2} [B, [B, A]], \\ S_{2,1} = [A + B, [A + B, C]], \\ S_{2,2} = -\frac{1}{2} [C, [C, A + B]]. \end{cases}$$

Therefore,

$$\frac{\text{Id} - P_{\Delta t}^{\gamma C, B, A, B, \gamma C}}{\Delta t} = -\mathcal{L}_\gamma - \frac{\Delta t}{2} \mathcal{L}_\gamma^2 - \Delta t^2 \left(\frac{1}{6} \mathcal{L}_\gamma^3 + S_2 \right) - \frac{\Delta t^3}{2} \left(\frac{1}{12} \mathcal{L}_\gamma^4 + S_2 \mathcal{L}_\gamma + \mathcal{L}_\gamma S_2 \right) - \Delta t^4 R_{2,\Delta t}. \quad (4.25)$$

We choose $f_2^{\gamma C, B, A, B, \gamma C}$ as the unique solution of the Poisson equation $\mathcal{L}_\gamma^* f_2^{\gamma C, B, A, B, \gamma C} = -S_2^* \mathbf{1}$ (which is indeed well posed since the right hand side has a vanishing average with respect to μ since it is in the image of S_2 , and is regular as shown by (4.26)). Then, for a function $\varphi \in \mathcal{S}$,

$$\begin{aligned} \int_{\mathcal{E}} \left(\frac{\text{Id} - P_{\Delta t}^{\gamma C, B, A, B, \gamma C}}{\Delta t} \right) \varphi \left(1 + \Delta t^2 f_2^{\gamma C, B, A, B, \gamma C} \right) d\mu = \\ -\frac{\Delta t^3}{2} \int_{\mathcal{E}} S_2 \mathcal{L}_\gamma \varphi + (\mathcal{L}_\gamma^2 \varphi) f_2^{\gamma C, B, A, B, \gamma C} d\mu - \Delta t^4 \int_{\mathcal{E}} [\tilde{R}_{2,\Delta t} \varphi + \hat{R}_{2,\Delta t} \varphi f_2^{\gamma C, B, A, B, \gamma C}] d\mu, \end{aligned}$$

where many terms cancel by the invariance of μ by $(\mathcal{L}_\gamma^\alpha)^*$ (for integer powers α). The leading order term on the right-hand side in fact vanishes since it can be rewritten as

$$\int_{\mathcal{E}} S_2 \mathcal{L}_\gamma \varphi + \mathcal{L}_\gamma^2 \varphi f_2^{\gamma C, B, A, B, \gamma C} d\mu = \int_{\mathcal{E}} \mathcal{L}_\gamma \varphi \left(S_2^* \mathbf{1} + \mathcal{L}_\gamma^* f_2^{\gamma C, B, A, B, \gamma C} \right) d\mu = 0.$$

Therefore,

$$\int_{\mathcal{E}} \left(\frac{\text{Id} - P_{\Delta t}^{\gamma C, B, A, B, \gamma C}}{\Delta t} \right) \varphi \left(1 + \Delta t^2 f_2^{\gamma C, B, A, B, \gamma C} \right) d\mu = -\Delta t^4 \int_{\mathcal{E}} [\tilde{R}_{2,\Delta t} \varphi + \hat{R}_{2,\Delta t} \varphi f_2^{\gamma C, B, A, B, \gamma C}] d\mu.$$

We then replace φ by $Q_{2,\Delta t} \psi$ where $Q_{2,\Delta t}$ is an approximate inverse satisfying

$$\frac{\text{Id} - P_{\Delta t}^{\gamma C, B, A, B, \gamma C}}{\Delta t} Q_{2,\Delta t} = \text{Id} + \Delta t^4 Z_{\Delta t}.$$

The proof is concluded as in Section 4.4.

PROOF OF (2.26). To evaluate the expression of $S_2^* \mathbf{1}$, we need to compute the actions of the adjoints of the various commutators. Using $C \mathbf{1} = (A + B) \mathbf{1} = 0$ and

$$C^* = C, \quad A^* = -A + g, \quad B^* = -B - g,$$

straightforward computations show that $S_{2,2}^* \mathbf{1} = S_{2,1}^* \mathbf{1} = 0$. In addition, since

$$A(g^2) = 2gAg, \quad B(g^2) = 2gBg,$$

it follows

$$\begin{aligned} ([A, [A, B]])^* \mathbf{1} &= (A^2 B - 2ABA + BA^2)^* \mathbf{1} = (B^* A^* - 2A^* B^* - (A^*)^2) g \\ &= ((B+g)(A-g) - 2(A-g)(B+g) - (A-g)^2) g \\ &= (BA - 2AB - A^2) g = -(A+B)Ag, \end{aligned}$$

where we have used $ABg = BAg$ (as can be checked by a direct computation). A similar computation shows that $([B, [B, A]])^* \mathbf{1} = (-AB + 2BA + B^2)g = (A+B)Bg = ABg$ (since $B^2g = 0$ by a direct verification). Finally,

$$S_2^* \mathbf{1} = -\frac{1}{12}(A+B) \left(A + \frac{B}{2} \right) g. \quad (4.26)$$

To obtain the expression of $f_2^{A,B,\gamma^C,B,A}$, we use the TU lemma with the operator

$$U_{\Delta t} = e^{\gamma \Delta t C/2} e^{\Delta t B/2} e^{\Delta t A/2}.$$

A simple computation shows that

$$U_{\Delta t}^* \mathbf{1} = \mathbf{1} + \frac{\Delta t^2}{8}(A+B)g + \Delta t^3 R_{\Delta t}^* \mathbf{1}.$$

In fact, it can be shown that the Δt^3 do not pollute the remainder since the next order correction in the invariant measure has to be of order Δt^4 (see (2.25)). The expressions of $f_2^{\gamma^C,A,B,A,\gamma^C}$ and $f_2^{B,A,\gamma^C,A,B}$ are obtained in a similar manner.

4.7 Proof of Corollary 2.2

The proof relies on the results of Theorem 2.7 and the TU lemma (Lemma 2.4). More precisely, the error estimate (2.27) is established by following the same lines of proof as for second order splitting schemes, except that the contributions of order Δt^3 do not vanish. We then use the TU lemma by considering the GLA evolution as the reference, and express the invariant measure of second order splitting schemes in terms of the invariant measure of the GLA scheme. For instance, consider $P_{\Delta t}^{\gamma^C,B,A,B}$ and $P_{\Delta t}^{\gamma^C,B,A,B,\gamma^C}$, in which case $U_{\Delta t} = e^{\gamma \Delta t C/2}$. Then,

$$\begin{aligned} \int_{\mathcal{E}} \psi d\mu_{\Delta t}^{\gamma^C,B,A,B,\gamma^C} &= \int_{\mathcal{E}} (U_{\Delta t} \psi) d\mu_{\Delta t}^{\gamma^C,B,A,B} \\ &= \int_{\mathcal{E}} U_{\Delta t} \psi d\mu + \Delta t^2 \int_{\mathcal{E}} (U_{\Delta t} \psi) f_2^{\gamma^C,B,A,B} d\mu + \Delta t^3 \int_{\mathcal{E}} (U_{\Delta t} \psi) f_3^{\gamma^C,B,A,B} d\mu + \Delta t^4 r_{\psi,\gamma,\Delta t} \\ &= \int_{\mathcal{E}} \psi d\mu + \Delta t^2 \int_{\mathcal{E}} \psi f_2^{\gamma^C,B,A,B} d\mu + \Delta t^3 \int_{\mathcal{E}} \psi \left(f_3^{\gamma^C,B,A,B} + \frac{\gamma}{2} C f_2^{\gamma^C,B,A,B} \right) d\mu + \Delta t^4 \tilde{r}_{\psi,\gamma,\Delta t}, \end{aligned}$$

where we have used the invariance of μ by $U_{\Delta t}$. The comparison with (2.25)-(2.26) gives the desired result.

4.8 Approximation of integrated correlation functions

Proof of Theorem 2.8. We first introduce the invariant measure for the numerical scheme, using the fact that $-\mathcal{L}^{-1}\psi$ has zero average with respect to μ :

$$\begin{aligned} \int_{\mathcal{E}} (-\mathcal{L}^{-1}\psi) \varphi d\mu &= \int_{\mathcal{E}} (-\mathcal{L}^{-1}\psi) \left(\varphi - \int_{\mathcal{E}} \varphi d\mu_{\Delta t} \right) d\mu \\ &= \int_{\mathcal{E}} (-\mathcal{L}^{-1}\psi) \left(\varphi - \int_{\mathcal{E}} \varphi d\mu_{\Delta t} \right) d\mu_{\Delta t} + \Delta t^\alpha r_{\Delta t}^{\psi,\varphi}, \end{aligned} \quad (4.27)$$

where $r_{\Delta t}^{\psi,\varphi}$ is uniformly bounded for Δt sufficiently small by (2.32). In addition, by (2.33),

$$\text{Id} = \left(\Delta t \sum_{n=0}^{+\infty} P_{\Delta t}^n \right) \left(\frac{\text{Id} - P_{\Delta t}}{\Delta t} \right) = - \left(\sum_{n=0}^{+\infty} P_{\Delta t}^n \right) \left(\mathcal{L} + \Delta t S_1 + \cdots + \Delta t^{\alpha-1} S_{\alpha-1} + \Delta t^\alpha \tilde{R}_{\alpha,\Delta t} \right),$$

so that

$$-\mathcal{L}^{-1} = \left(\Delta t \sum_{n=0}^{+\infty} P_{\Delta t}^n \right) \left(\text{Id} + \cdots + \Delta t^{\alpha-1} S_{\alpha-1} \mathcal{L}^{-1} + \Delta t^\alpha \tilde{R}_{\alpha, \Delta t} \mathcal{L}^{-1} \right). \quad (4.28)$$

Note finally that, for any $n \geq 0$ and $\phi \in \mathcal{S}$,

$$\begin{aligned} \int_{\mathcal{E}} P_{\Delta t}^n \phi \left(\varphi - \int_{\mathcal{E}} \varphi d\mu_{\Delta t} \right) d\mu_{\Delta t} &= \int_{\mathcal{E}} P_{\Delta t}^n \left(\phi - \int_{\mathcal{E}} \phi d\mu_{\Delta t} \right) \left(\varphi - \int_{\mathcal{E}} \varphi d\mu_{\Delta t} \right) d\mu_{\Delta t} \\ &= \int_{\mathcal{E}} P_{\Delta t}^n \left(\phi - \int_{\mathcal{E}} \phi d\mu_{\Delta t} \right) \varphi d\mu_{\Delta t}. \end{aligned}$$

We now replace ϕ by $\Phi_{\Delta t, \alpha} = \psi_{\Delta t, \alpha} + \Delta t^\alpha \tilde{R}_{\Delta t, \alpha} \mathcal{L}^{-1} \psi$. From (4.27)-(4.28), the series appearing on the right-hand side of the inequality below is well defined:

$$\begin{aligned} \int_{\mathcal{E}} (-\mathcal{L}^{-1} \psi) \left(\varphi - \int_{\mathcal{E}} \varphi d\mu_{\Delta t} \right) d\mu_{\Delta t} &= \Delta t \int_{\mathcal{E}} \sum_{n=0}^{+\infty} P_{\Delta t}^n \left(\Phi_{\Delta t, \alpha} - \int_{\mathcal{E}} \Phi_{\Delta t, \alpha} d\mu_{\Delta t} \right) \left(\varphi - \int_{\mathcal{E}} \varphi d\mu_{\Delta t} \right) d\mu_{\Delta t} \\ &= \Delta t \int_{\mathcal{E}} \sum_{n=0}^{+\infty} P_{\Delta t}^n \left(\Phi_{\Delta t, \alpha} - \int_{\mathcal{E}} \Phi_{\Delta t, \alpha} d\mu_{\Delta t} \right) \varphi d\mu_{\Delta t}. \end{aligned} \quad (4.29)$$

It remains to bound the extra term arising from $\Phi_{\Delta t, \alpha} - \psi_{\Delta t, \alpha} = \Delta t^\alpha \tilde{R}_{\Delta t, \alpha} \mathcal{L}^{-1} \psi$. We use to this end Corollary 2.1: introducing

$$\Psi_{\Delta t, \alpha} = \tilde{R}_{\Delta t, \alpha} \mathcal{L}^{-1} \psi - \int_{\mathcal{E}} \tilde{R}_{\Delta t, \alpha} \mathcal{L}^{-1} \psi d\mu_{\Delta t},$$

it holds, uniformly in Δt (for Δt small enough)

$$\left| \Delta t \int_{\mathcal{E}} \sum_{n=0}^{+\infty} (P_{\Delta t}^n \Psi_{\Delta t, \alpha}) \varphi d\mu_{\Delta t} \right| = \left| \int_{\mathcal{E}} \left(\frac{\text{Id} - P_{\Delta t}}{\Delta t} \right)^{-1} \Psi_{\Delta t, \alpha} \varphi d\mu_{\Delta t} \right| \leq R.$$

The combination of this bound and (4.27)-(4.29) gives (2.35). \square

Proof of Corollary 2.3. Comparing (2.37) and (2.33), we see that $S_1 = \mathcal{L}^2/2$. The idea is to start from (2.35) and to appropriately rewrite the first order correction term. We use to this end (2.35) with ψ replaced by its first order correction $(\psi_{\Delta t, 2} - \psi)/\Delta t = S_1 \mathcal{L}^{-1} \psi$, and discard terms of order Δt^2 :

$$\int_0^{+\infty} \mathbb{E} \left(S_1 \mathcal{L}^{-1} \psi(q_t, p_t) \varphi(q_0, p_0) \right) dt = \Delta t \sum_{n=0}^{+\infty} \mathbb{E}_{\Delta t} \left(S_1 \mathcal{L}^{-1} \psi(q^{n+1}, p^{n+1}) \varphi_{\Delta t, 0}(q^0, p^0) \right) + \Delta t r_{\Delta t}^{\psi, \varphi},$$

where $r_{\Delta t}^{\psi, \varphi}$ is uniformly bounded for Δt sufficiently small and

$$\varphi_{\Delta t, 0} = \varphi - \int_{\mathcal{E}} \varphi d\mu_{\Delta t}.$$

On the other hand,

$$\int_0^{+\infty} \mathbb{E} \left(S_1 \mathcal{L}^{-1} \psi(q_t, p_t) \varphi(q_0, p_0) \right) dt = - \int_{\mathcal{E}} \mathcal{L}^{-1} S_1 \mathcal{L}^{-1} \psi \varphi d\mu = - \frac{1}{2} \int_{\mathcal{E}} \psi \varphi d\mu,$$

so that

$$\begin{aligned} &\Delta t \sum_{n=0}^{+\infty} \mathbb{E}_{\Delta t} \left((S_1 \mathcal{L}^{-1} \psi)_{\Delta t, 0}(q^{n+1}, p^{n+1}) \varphi(q^0, p^0) \right) \\ &= \Delta t \sum_{n=0}^{+\infty} \mathbb{E}_{\Delta t} \left(S_1 \mathcal{L}^{-1} \psi(q^{n+1}, p^{n+1}) \varphi_{\Delta t, 0}(q^0, p^0) \right) \\ &= \int_0^{+\infty} \mathbb{E} \left(S_1 \mathcal{L}^{-1} \psi(q_t, p_t) \varphi_{\Delta t, 0}(q_0, p_0) \right) dt - \Delta t r_{\Delta t}^{\psi, \varphi} \\ &= - \frac{1}{2} \int_{\mathcal{E}} \psi \varphi_{\Delta t, 0} d\mu - \Delta t r_{\Delta t}^{\psi, \varphi} = - \frac{1}{2} \int_{\mathcal{E}} \psi_{\Delta t, 0} \varphi d\mu - \Delta t \tilde{r}_{\Delta t}^{\psi, \varphi} \\ &= - \frac{1}{2} \mathbb{E}_{\Delta t} (\psi_{\Delta t, 0} \varphi) + \Delta t \tilde{r}_{\Delta t}^{\psi, \varphi}. \end{aligned}$$

This gives (2.38). \square

4.9 Proof of Theorem 2.9

We write the proof for the evolution operator $P_{\Delta t}^{\gamma C, A, B, A, \gamma C}$ first, and mention then how to extend the result to $P_{\Delta t}^{B, A, \gamma C, A, B}$ using the TU lemma. The proofs for $P_{\Delta t}^{\gamma C, B, A, B, \gamma C}$ and $P_{\Delta t}^{A, B, \gamma C, B, A}$ are very similar, so we only briefly mention the required modifications.

REDUCTION TO A LIMITING OPERATOR UP TO EXPONENTIALLY SMALL TERMS. Let us introduce the evolution operator corresponding to the standard position Verlet scheme: $P_{\text{ham}, \Delta t} = e^{\Delta t A/2} e^{\Delta t B} e^{\Delta t A/2}$, so that $P_{\Delta t}^{\gamma C, A, B, A, \gamma C} = e^{\gamma \Delta t C/2} P_{\text{ham}, \Delta t} e^{\gamma \Delta t C/2}$. On the other hand, we have the following convergence result, whose proof is postponed to the end of the section.

LEMMA 4.2 Fix $s^* \in \mathbb{N}^*$. Then, there exist $K, \alpha > 0$ such that, for any $1 \leq s \leq s^*$ and any $t \geq 0$,

$$\|e^{\gamma t C} - \pi\|_{\mathcal{B}(L^\infty_{\mathcal{X}_s})} \leq K e^{-\alpha \gamma t}.$$

This suggests to consider the limiting operator $P_{\infty, \Delta t} = \pi P_{\text{ham}, \Delta t} \pi$ and to write

$$P_{\Delta t}^{\gamma C, A, B, A, \gamma C} - P_{\infty, \Delta t} = (e^{\gamma \Delta t C/2} - \pi) P_{\text{ham}, \Delta t} \pi + e^{\gamma \Delta t C/2} P_{\text{ham}, \Delta t} (e^{\gamma \Delta t C/2} - \pi). \quad (4.30)$$

For a given smooth function $\varphi \in \mathcal{S}$ which depends only on the position variable q ,

$$\int_{\mathcal{E}} (\text{Id} - P_{\Delta t}^{\gamma C, A, B, A, \gamma C}) \varphi d\mu_{\gamma, \Delta t} = 0 = \int_{\mathcal{E}} (\text{Id} - P_{\infty, \Delta t}) \varphi d\mu_{\gamma, \Delta t} + r_{\varphi, \gamma, \Delta t}^1, \quad (4.31)$$

with the remainder

$$r_{\varphi, \gamma, \Delta t}^1 = \int_{\mathcal{E}} (P_{\infty, \Delta t} - P_{\Delta t}^{\gamma C, A, B, A, \gamma C}) \varphi d\mu_{\gamma, \Delta t}.$$

On the other hand,

$$\int_{\mathcal{E}} [(\text{Id} - P_{\Delta t}^{\gamma C, B, A, B, \gamma C}) \varphi] (1 + \Delta t^2 f_{2, \infty}) d\mu = \int_{\mathcal{E}} [(\text{Id} - P_{\infty, \Delta t}) \varphi] (1 + \Delta t^2 f_{2, \infty}) d\mu + r_{\varphi, \gamma, \Delta t}^2, \quad (4.32)$$

with the remainder

$$r_{\varphi, \gamma, \Delta t}^2 = \int_{\mathcal{E}} [(P_{\infty, \Delta t} - P_{\Delta t}^{\gamma C, B, A, B, \gamma C}) \varphi] (1 + \Delta t^2 f_{2, \infty}) d\mu.$$

The idea is that the remainders are exponentially small when the function φ is sufficiently smooth (see below for a more precise discussion, once φ has been replaced by $Q_{\Delta t} \psi$ with $Q_{\Delta t}$ an appropriate approximate inverse). Therefore, the leading order terms in the error estimate are obtained by considering the limiting operator $P_{\infty, \Delta t}$ only.

ERROR ESTIMATES FOR THE LIMITING OPERATOR $P_{\infty, \Delta t}$. We now study the error estimates associated with $P_{\infty, \Delta t}$, following the strategy used in Section 4.4. We first use the results of Section 4.3.2 with $M = 3$, $A_1 = A_3 = A/2$ and $A_2 = B$ to expand $P_{\text{ham}, \Delta t}$ as

$$P_{\infty, \Delta t} = \pi + \Delta t \pi (A + B) \pi + \frac{\Delta t^2}{2} \pi (A + B)^2 \pi + \frac{\Delta t^3}{6} \pi S_3 \pi + \frac{\Delta t^4}{24} \pi S_4 \pi + \frac{\Delta t^5}{120} \pi S_5 \pi + \Delta t^6 \pi R_{\Delta t} \pi, \quad (4.33)$$

with $S_i = \mathcal{S}[(A_1 + A_2 + A_3)^i]$. To give more precise expressions of the operators appearing on the right-hand side of the above equality, we use the following facts:

$$\forall n \in \mathbb{N}, \quad B^n \pi = 0, \quad \pi A^{2n+1} \pi = 0, \quad (4.34)$$

and

$$\forall n \geq m + 1, \quad B^n A^m \pi = 0. \quad (4.35)$$

In addition,

$$\pi A^2 \pi = \frac{1}{\beta} \Delta_q \pi, \quad B A \pi = -\nabla V \cdot \nabla_q \pi.$$

Using these rules in (4.33) leads to

$$\pi(A+B)\pi = 0, \quad \pi(A+B)^2\pi = \pi(A^2+BA)\pi = \mathcal{L}_{\text{ovd}}\pi. \quad (4.36)$$

The operator S_3 is a combination of terms of the form $A^a B^b A^c$ with $a+b+c=3$ and $a, b, c \in \mathbb{N}$. In view of (4.34)-(4.35), only the terms with $c \geq 1$ and $b \leq c$ have to be considered, so that only BA^2 and ABA remain. A simple computation shows that $BA^2\pi\varphi$ and $ABA\pi\varphi$ are functions linear in p , so that $\pi BA^2\pi = \pi ABA\pi\varphi = 0$. Finally, $\pi S_3\pi = 0$. A similar reasoning shows that $\pi S_5\pi = 0$ and that many terms appearing in the expression of S_4 also disappear.

Plugging the above results in (4.33) and introducing $h = \Delta t^2/2$,

$$P_{\infty, \Delta t} = \pi + h\pi\mathcal{L}_{\text{ovd}}\pi + \frac{h^2}{6}\pi\left(A^4 + \frac{3}{2}A^2BA + \frac{3}{2}ABA^2 + \frac{3}{2}B^2A^2 + \frac{1}{2}BA^3\right)\pi + h^3R_{\infty, \Delta t}.$$

Using

$$\begin{aligned} \pi A^4\pi\varphi &= \frac{3}{\beta^2}\Delta_q^2\pi\varphi = 3(\pi A^2\pi)^2\varphi, \\ \pi BA^3\pi\varphi &= -\frac{3}{\beta}\nabla V \cdot \nabla_q(\Delta_q\pi\varphi) = 3\pi BA\pi A^2\pi\varphi, \\ \pi B^2A^2\pi\varphi &= 2(\nabla V)^T(\nabla_q^2\pi\varphi)\nabla V, \\ \pi ABA^2\pi\varphi &= -\frac{2}{\beta}(\nabla^2 V : \nabla^2\varphi + \nabla V \cdot \nabla(\Delta\varphi)), \\ \pi A^2BA\pi\varphi &= -\frac{1}{\beta}(2\nabla^2 V : \nabla^2\varphi + \nabla V \cdot \nabla(\Delta\varphi) + \nabla(\Delta V) \cdot \nabla\varphi) = \pi A^2\pi BA\pi, \end{aligned} \quad (4.37)$$

it follows

$$\begin{aligned} &\left(A^4 + \frac{3}{2}A^2BA + \frac{3}{2}ABA^2 + \frac{3}{2}B^2A^2 + \frac{1}{2}BA^3\right)\pi\varphi \\ &= \frac{3}{\beta^2}\Delta_q^2\varphi - \frac{6}{\beta}\nabla^2 V : \nabla^2\varphi - \frac{6}{\beta}\nabla V \cdot \nabla(\Delta\varphi) - \frac{3}{2\beta}\nabla(\Delta V) \cdot \nabla\varphi + 3(\nabla V)^T(\nabla^2\varphi)\nabla V. \end{aligned}$$

A straightforward computation shows that

$$\mathcal{L}_{\text{ovd}}^2 = \frac{1}{\beta^2}\Delta_q^2 - \frac{2}{\beta}\nabla^2 V : \nabla^2\varphi - \frac{2}{\beta}\nabla V \cdot \nabla(\Delta\varphi) - \frac{1}{\beta}\nabla(\Delta V) \cdot \nabla\varphi + (\nabla V)^T(\nabla^2\varphi)\nabla V + (\nabla V)^T(\nabla^2 V)\nabla\varphi.$$

Therefore,

$$\pi\left(A^4 + \frac{3}{2}A^2BA + \frac{3}{2}ABA^2 + \frac{3}{2}B^2A^2 + \frac{1}{2}BA^3\right)\pi = 3(\mathcal{L}_{\text{ovd}}^2 + D)\pi,$$

with

$$D\varphi = \frac{1}{2\beta}\nabla(\Delta V) \cdot \nabla\varphi - (\nabla V)^T(\nabla^2 V)\nabla\varphi. \quad (4.38)$$

In conclusion,

$$P_{\infty, \Delta t} = \pi + h\mathcal{L}_{\text{ovd}}\pi + \frac{h^2}{2}(\mathcal{L}_{\text{ovd}}^2 + D)\pi + h^3R_{\infty, \Delta t}. \quad (4.39)$$

Let us emphasize that this operator acts on functions of q , that π is the identity operator for functions which do not depend on p , and note that

$$\frac{\pi - P_{\infty, \Delta t}}{h} = -\mathcal{L}_{\text{ovd}} - \frac{h}{2}(\mathcal{L}_{\text{ovd}}^2 + D) - h^2R_{\Delta t}. \quad (4.40)$$

An approximate inverse of (4.40) is therefore

$$Q_h = -\mathcal{L}_{\text{ovd}}^{-1} + \frac{h}{2}(\text{Id} + \mathcal{L}_{\text{ovd}}^{-1}D\mathcal{L}_{\text{ovd}}^{-1}).$$

Denote by $\bar{\mu}_{\infty, \Delta t}(\mathrm{d}q)$ the invariant measure of the Markov chain generated by the limiting method $P_{\infty, \Delta t}$. Proceeding as in Section 4.4 by first identifying the leading order correction $f_{2, \infty}$ and replacing φ by $Q_h \psi$, the equality (4.39) allows us to show that

$$\int_{\mathcal{M}} \psi(q) \bar{\mu}_{\infty, \Delta t}(\mathrm{d}q) = \int_{\mathcal{M}} \psi(q) \bar{\mu}(\mathrm{d}q) + \Delta t^2 \int_{\mathcal{M}} \psi(q) f_{2, \infty}(q) \bar{\mu}(\mathrm{d}q) + \Delta t^4 \bar{r}_{\Delta t, \psi}, \quad (4.41)$$

where $f_{2, \infty}$ is the unique solution of

$$\mathcal{L}_{\text{ovd}} f_{2, \infty} = -\frac{1}{4} D^* \mathbf{1}. \quad (4.42)$$

A more explicit expression can be obtained by noting that

$$D\varphi = \frac{1}{2} \nabla \left(\frac{1}{\beta} \Delta V - |\nabla V|^2 \right) \cdot \nabla \varphi,$$

so that (recalling $\mathcal{L}_{\text{ovd}} = -\beta^{-1} \nabla^* \nabla = -\beta^{-1} \sum_{i=1}^{dN} \partial_{q_i}^* \partial_{q_i}$ where adjoints are taken on $L^2(\bar{\mu})$)

$$\begin{aligned} \int_{\mathcal{M}} \varphi (D^* \mathbf{1}) \, \mathrm{d}\bar{\mu} &= \int_{\mathcal{M}} D\varphi \, \mathrm{d}\bar{\mu} = \frac{1}{2} \int_{\mathcal{M}} \varphi \nabla^* \nabla \left(\frac{1}{\beta} \Delta V - |\nabla V|^2 \right) \, \mathrm{d}\bar{\mu} \\ &= -\frac{1}{2} \int_{\mathcal{M}} \varphi \mathcal{L}_{\text{ovd}} (\Delta V - \beta |\nabla V|^2) \, \mathrm{d}\bar{\mu}. \end{aligned}$$

Since $f_{2, \infty}$ should have a vanishing average with respect to μ , this proves that

$$f_{2, \infty}(q) = \frac{1}{8} (\Delta V - \beta |\nabla V|^2) + a, \quad (4.43)$$

where the constant a is adjusted to account for the constraint of vanishing average. A simple computation shows that it is equal to the constant $a_{\beta, V}$ defined in (2.41).

In fact, it is possible for the scheme considered here to have an even more explicit expression of the leading order correction term to numerical averages by noting that

$$\frac{1}{\beta} \int_{\mathcal{M}} \Delta \varphi \, \mathrm{d}\bar{\mu} = - \int_{\mathcal{M}} \varphi (\Delta V - \beta |\nabla V|^2) \, \mathrm{d}\bar{\mu}, \quad (4.44)$$

so that finally

$$\int_{\mathcal{M}} \psi(q) \bar{\mu}_{\infty, \Delta t}(\mathrm{d}q) = \int_{\mathcal{M}} \psi(q) \bar{\mu}(\mathrm{d}q) - \frac{\Delta t^2}{8\beta} \int_{\mathcal{M}} \Delta \psi(q) \bar{\mu}(\mathrm{d}q) + \Delta t^4 \bar{r}_{\Delta t, \psi}.$$

CONCLUSION OF THE PROOF. We now come back to (4.31)-(4.32) and replace φ by $Q_h \psi$:

$$\int_{\mathcal{E}} \psi \, \mathrm{d}\mu_{\gamma, \Delta t} = \int_{\mathcal{E}} \psi (1 + \Delta t^2 f_{2, \infty}) \, \mathrm{d}\mu + r_{\psi, \gamma, \Delta t}^1 + r_{\psi, \gamma, \Delta t}^2 + \Delta t^4 \bar{r}_{\Delta t, \psi}, \quad (4.45)$$

where $\bar{r}_{\Delta t, \psi}$ is the same as in (4.41), while

$$\begin{aligned} r_{\psi, \gamma, \Delta t}^1 &= \int_{\mathcal{E}} \left(P_{\infty, \Delta t} - P_{\Delta t}^{\gamma C, A, B, A, \gamma C} \right) Q_h \psi \, \mathrm{d}\mu_{\gamma, \Delta t}, \\ r_{\psi, \gamma, \Delta t}^2 &= \int_{\mathcal{E}} \left[\left(P_{\infty, \Delta t} - P_{\Delta t}^{\gamma C, B, A, B, \gamma C} \right) Q_h \psi \right] (1 + \Delta t^2 f_{2, \infty}) \, \mathrm{d}\mu. \end{aligned}$$

We then integrate with respect to momenta in (4.45), and bound the remainders by $Ke^{-\kappa \gamma \Delta t}$ in view of the decomposition (4.30) and Lemma 4.2 (the operators $P_{\text{ham}, \Delta t}$ and $e^{\gamma \Delta t C/2}$ being bounded on $L_{\mathcal{H}_s}^\infty$ uniformly in Δt).

PROOF OF (2.41) FOR $f_{2,\infty}^{B,A,\gamma^C,A,B}$ We set

$$U_{\gamma,\Delta t} = e^{\gamma\Delta t C/2} e^{\Delta t A/2} e^{\Delta t B/2}, \quad T_{\gamma,\Delta t} = e^{\Delta t B/2} e^{\Delta t A/2} e^{\gamma\Delta t C/2},$$

so that $P_{\Delta t}^{B,A,\gamma^C,A,B} = T_{\gamma,\Delta t} U_{\gamma,\Delta t}$ while $P_{\Delta t}^{\gamma^C,A,B,A,\gamma^C} = U_{\gamma,\Delta t} T_{\gamma,\Delta t}$. By the TU lemma,

$$\begin{aligned} \int_{\mathcal{E}} \psi d\mu_{\Delta t}^{B,A,\gamma^C,A,B} &= \int_{\mathcal{E}} (U_{\gamma,\Delta t} \psi) d\mu_{\Delta t}^{\gamma^C,A,B,A,\gamma^C} \\ &= \int_{\mathcal{E}} (U_{\infty,\Delta t} \psi) d\mu_{\Delta t}^{\gamma^C,A,B,A,\gamma^C} + \int_{\mathcal{E}} (U_{\gamma,\Delta t} - U_{\infty,\Delta t}) \psi d\mu_{\Delta t}^{\gamma^C,A,B,A,\gamma^C}, \end{aligned}$$

where we have introduced $U_{\infty,\Delta t} = \pi e^{\Delta t A/2} e^{\Delta t B/2}$. The second term can be bounded $Ke^{-\kappa\gamma\Delta t}$ in view of Lemma 4.2 and the moment estimate (2.17). For the first one, we use (4.45) and the following expansion (using the rules (4.34)-(4.35)):

$$U_{\infty,\Delta t} \psi = U_{\infty,\Delta t} \pi \psi = \psi + \frac{\Delta t^2}{8} \pi A^2 \pi \psi + \Delta t^4 \tilde{r}_{\psi,\Delta t} = \psi + \frac{\Delta t^2}{8\beta} \Delta \psi + \Delta t^4 \tilde{r}_{\psi,\Delta t},$$

where the remainder $\tilde{r}_{\psi,\Delta t}$ is uniformly bounded for Δt sufficiently small. Therefore,

$$\int_{\mathcal{E}} (U_{\infty,\Delta t} \psi) d\mu_{\Delta t}^{\gamma^C,A,B,A,\gamma^C} = \int_{\mathcal{E}} \psi (1 + \Delta t^2 f_{2,\infty}) d\mu + \frac{\Delta t^2}{8\beta} \int_{\mathcal{E}} \Delta \psi d\mu + \hat{r}_{\psi,\gamma,\Delta t},$$

where $f_{2,\infty}$ is given in (4.43). The remainder $\hat{r}_{\psi,\gamma,\Delta t}$ is the sum of terms of order Δt^4 and others which can be bounded by $Ke^{-\kappa\gamma\Delta t}$. We conclude by resorting to (4.44) to compute the adjoint of the operator Δ_q on $L^2(\mu)$.

PROOF OF (2.41) FOR $f_{2,\infty}^{\gamma^C,B,A,B,\gamma^C}$ AND $f_{2,\infty}^{A,B,\gamma^C,B,A}$. We mimic the above proof for the evolution operator $P_{\Delta t}^{\gamma^C,B,A,B,\gamma^C}$. The equality (4.33) still holds, but the operator S_4 now reads

$$S_4 = A^4 + 2BA^2 + \frac{3}{2}B^2A^2,$$

so that

$$D\phi = \frac{2}{\beta} \nabla^2 V : \nabla^2 \phi + \frac{1}{\beta} \nabla(\Delta V) \cdot \nabla \phi - \nabla V^T (\nabla^2 V) \nabla \phi.$$

A simple computation shows that

$$\int_{\mathcal{M}} D\phi d\bar{\mu} = -\frac{1}{\beta} \int_{\mathcal{M}} \nabla \left(\Delta V - \frac{\beta}{2} |\nabla V|^2 \right) \cdot \nabla \phi d\bar{\mu} = \int_{\mathcal{M}} \mathcal{L}_{\text{ovd}} \left(\Delta V - \frac{\beta}{2} |\nabla V|^2 \right) \phi d\bar{\mu},$$

so that, in view of (4.42),

$$f_{2,\infty}^{\gamma^C,B,A,B,\gamma^C} = -\frac{1}{4} \left(\Delta V - \frac{\beta}{2} |\nabla V|^2 \right).$$

The expression of $f_{2,\infty}^{A,B,\gamma^C,B,A}$ is obtained via the TU lemma, introducing the limiting operator

$$U_{\infty,\Delta t} \pi = \pi e^{\Delta t B/2} e^{\Delta t A/2} \pi = \pi + \frac{\Delta t^2}{8} \pi (A^2 + 2BA) \pi + \Delta t^4 R_{\Delta t},$$

so that

$$f_{2,\infty}^{A,B,\gamma^C,B,A} = f_{2,\infty}^{\gamma^C,B,A,B,\gamma^C} + \frac{1}{8} \left(\pi (A^2 + 2BA) \pi \right)^* \mathbf{1} = f_{2,\infty}^{\gamma^C,B,A,B,\gamma^C} + \frac{1}{8} \left(\pi BA \pi \right)^* \mathbf{1} = -\frac{1}{8} \Delta V.$$

Let us conclude this section with the

Proof of Lemma 4.2. The conclusion follows for instance by an application of (Rey-Bellet, 2006, Theorem 8.7), considering as a reference dynamics the Ornstein-Uhlenbeck process

$$dp_t = -M^{-1}p_t dt + \sqrt{\frac{2\gamma}{\beta}} dW_t$$

with generator C (recall that the unique invariant probability measure of this process is $\kappa(dp)$). To apply the theorem, we need to show that \mathcal{K}_s is a Lyapunov function for any $s \geq 1$. We compute

$$C\mathcal{K}_s = \left(-2sp^T p + \frac{2s(dN + 2s - 2)}{\beta} \right) |p|^{2(s-1)} \leq -\mathcal{K}_s + b_s$$

for an appropriate constant $b_s \geq 0$. This shows the existence of constants R_s, α_s such that

$$\left| (e^{tC} f)(p) - \int_{\mathbb{R}^{dN}} f(p) \kappa(dp) \right| \leq R_s e^{-\alpha_s t} \|f\|_{L_{\mathcal{K}_s}^\infty(dp)} \mathcal{K}_s(p),$$

where the notation $L_{\mathcal{K}_s}^\infty(dp)$ emphasizes that the supremum is taken over a function of the momentum variable only. The desired result now follows by applying the above bound to the function $\psi(q, \cdot)$ for any element $\psi \in L_{\mathcal{K}_s}^\infty$, and taking the supremum over q . \square

4.10 Proof of Proposition 2.10

Recall that we set $M = \text{Id}$ for overdamped limits. We consider first $f_2^{\gamma^C, B, A, B, \gamma^C}$, which satisfies (2.26). Let us first compute the right-hand side. Since

$$\left[\left(A + \frac{1}{2}B \right) g \right] = \beta \left(p^T (\nabla^2 V) p - \frac{1}{2} |\nabla V|^2 \right),$$

a simple computation shows that

$$\tilde{g} = \frac{1}{12} (A + B) \left[\left(A + \frac{1}{2}B \right) g \right] = \frac{\beta}{12} \left[(\nabla^3 V) : (p \otimes p \otimes p) - 3p^T (\nabla^2 V) \nabla V \right].$$

Note that the above function has average 0 with respect to κ . We then apply Theorem 2.2 to obtain

$$\left\| f_2^{\gamma^C, B, A, B, \gamma^C} - \mathcal{L}_{\text{ovd}}^{-1} \pi(A + B) C^{-1} \tilde{g} \right\|_{H^1(\mu)} \leq \frac{K}{\gamma}.$$

Since

$$C \left[(\nabla^3 V) : (p \otimes p \otimes p) \right] = -3(\nabla^3 V) : (p \otimes p \otimes p) + \frac{6}{\beta} p^T \nabla (\Delta V),$$

it is easily checked that

$$\begin{aligned} C^{-1} \tilde{g} &= -\frac{\beta}{36} (\nabla^3 V) : (p \otimes p \otimes p) - \frac{1}{6} p^T \nabla (\Delta V) + \frac{\beta}{4} p^T (\nabla^2 V) \nabla V \\ &= -\frac{\beta}{36} A^3 \pi V + A \pi \left(-\frac{1}{6} (\Delta V) + \frac{\beta}{8} |\nabla V|^2 \right). \end{aligned}$$

To compute $\pi(A + B) C^{-1} \tilde{g}$, we rely on (4.36) and (4.37) and obtain

$$\begin{aligned} \pi(A + B) C^{-1} \tilde{g} &= -\frac{1}{12} \left(\frac{1}{\beta} \Delta^2 V - \nabla V \cdot \nabla (\Delta V) \right) + \mathcal{L}_{\text{ovd}} \left(-\frac{1}{6} (\Delta V) + \frac{\beta}{8} |\nabla V|^2 \right) \\ &= \mathcal{L}_{\text{ovd}} \left(-\frac{1}{4} \Delta V + \frac{\beta}{8} |\nabla V|^2 \right). \end{aligned}$$

This allows us to conclude that the limit of $f_2^{\gamma^C, B, A, B, \gamma^C}$ is the argument of the operator \mathcal{L}_{ovd} in the previous line, up to an additive constant chosen to ensure that $f_2^{\gamma^C, B, A, B, \gamma^C}$ has a vanishing average with respect to μ (which turns out to be $a_{\beta, V}$). We deduce the limit for $f_2^{A, B, \gamma^C, B, A}$ with (2.26) since $(A + B)g = p^T (\nabla^2 V) p - |\nabla V|^2$.

The expressions for the limits of $f_2^{\gamma^C, A, B, A, \gamma^C}$ and $f_2^{B, A, \gamma^C, A, B}$ are obtained in a similar fashion.

4.11 Linear response theory

4.11.1 *Definition of the mobility* (3.4). We briefly sketch the discussion in (Stoltz, 2012, Section 3.1) (see in particular Theorem 3.1 in this reference). Hypocoercivity arguments show that the measure $\mu_{\gamma,\eta}$ has a smooth density with respect to the Lebesgue measure. It moreover formally satisfies the Fokker-Planck equation

$$\left(\mathcal{L}_\gamma + \eta \widetilde{\mathcal{L}}\right)^* f_{\gamma,\eta} = 0, \quad \mu_{\gamma,\eta}(\mathrm{d}q \mathrm{d}p) = f_{\gamma,\eta}(q, p) \mu(\mathrm{d}q \mathrm{d}p), \quad \int_{\mathcal{E}} \mathrm{d}\mu_{\gamma,\eta} = 1. \quad (4.46)$$

This equation can be given a rigorous meaning when $f_{\gamma,\eta} \in L^2(\mu)$, which is the case when η is sufficiently small. We rely on the following result (proved at the end of this section).

LEMMA 4.3 The operator $(\mathcal{L}_\gamma^*)^{-1} \widetilde{\mathcal{L}}^*$, considered as an operator on the Hilbert space $\mathcal{H}^0 = L^2(\mu) \cap \{\mathbf{1}\}^\perp$ introduced in (2.7), is bounded.

Denoting by r the spectral radius of $(\mathcal{L}_\gamma^*)^{-1} \widetilde{\mathcal{L}}^* \in \mathcal{B}(\mathcal{H}^0)$, it is easily checked that $(\mathcal{L}_\gamma + \eta \widetilde{\mathcal{L}})^*$ is invertible for $|\eta| < r^{-1}$ with

$$\left[(\mathcal{L}_\gamma + \eta \widetilde{\mathcal{L}})^*\right]^{-1} = \left(\sum_{n=0}^{+\infty} (-\eta)^n \left[(\mathcal{L}_\gamma^*)^{-1} \widetilde{\mathcal{L}}^*\right]^n\right) (\mathcal{L}_\gamma^*)^{-1}.$$

Therefore, a straightforward computation shows that

$$f_{\gamma,\eta}(q, p) = 1 + \sum_{n=1}^{+\infty} (-\eta)^n \left[(\mathcal{L}_\gamma^*)^{-1} \widetilde{\mathcal{L}}^*\right]^n \mathbf{1} \quad (4.47)$$

is an admissible solution of (4.46), and it is in fact the only one in view of the uniqueness of the invariant probability measure. Note that the normalization of the measure $f_{\gamma,\eta} \mathrm{d}\mu$ does not depend on η . Finally,

$$\int_{\mathcal{E}} F^T M^{-1} p \mu_{\gamma,\eta}(\mathrm{d}q \mathrm{d}p) = -\eta \int_{\mathcal{E}} F^T M^{-1} p \left[(\mathcal{L}_\gamma^*)^{-1} \widetilde{\mathcal{L}}^* \mathbf{1}\right] \mu(\mathrm{d}q \mathrm{d}p) + \eta^2 r_{\eta,\gamma},$$

with $r_{\eta,\gamma}$ uniformly bounded as $\eta \rightarrow 0$. This gives (3.4).

Proof of Lemma 4.3. Note first that the image of $\widetilde{\mathcal{L}}^*$ is contained in \mathcal{H}^0 since, for any $u \in \mathcal{H}^1$ (the Hilbert space defined in (2.6)),

$$\int_{\mathcal{E}} \widetilde{\mathcal{L}}^* u \mathrm{d}\mu = \int_{\mathcal{E}} u \left(\widetilde{\mathcal{L}} \mathbf{1}\right) \mathrm{d}\mu = 0.$$

It is therefore possible to give a meaning to the operator $(\mathcal{L}_\gamma^*)^{-1} \widetilde{\mathcal{L}}^*$. We then check that the perturbation $\widetilde{\mathcal{L}}$ is \mathcal{L}_γ -bounded (with relative bound 0, in fact):

$$\left\|\widetilde{\mathcal{L}} u\right\|_{L^2(\mu)}^2 \leq |F|^2 \|\nabla_p u\|_{L^2(\mu)}^2 = -\beta |F|^2 \langle u, \mathcal{L}_\gamma u \rangle_{L^2(\mu)} \leq \beta |F|^2 \|u\|_{L^2(\mu)} \|\mathcal{L}_\gamma u\|_{L^2(\mu)},$$

so that, for $u \in \mathcal{H}^0$ (recall that $\mathcal{L}_\gamma^{-1} u$ is well defined in this case),

$$\left\|\widetilde{\mathcal{L}} \mathcal{L}_\gamma^{-1} u\right\|_{L^2(\mu)}^2 \leq \beta |F|^2 \|u\|_{L^2(\mu)} \|\mathcal{L}_\gamma^{-1} u\|_{L^2(\mu)} \leq \beta |F|^2 \|\mathcal{L}_\gamma^{-1}\|_{\mathcal{B}(\mathcal{H}^0)} \|u\|_{L^2(\mu)}^2.$$

This proves that $\widetilde{\mathcal{L}} \mathcal{L}_\gamma^{-1}$ is bounded, hence its adjoint is bounded as well. \square

4.11.2 *Proof of Lemma 3.1.* Recall that we set mass matrices to identity when considering overdamped limits. Since

$$\mathcal{L}_\gamma(F^T p) = -\gamma F^T p - F^T \nabla V,$$

it follows (using first (4.47) to compute the linear response and then (2.10) to obtain the behavior of $\mathcal{L}_\gamma^{-1}(F^T \nabla V)$)

$$\begin{aligned}
\gamma v_{F,\gamma} &= \lim_{\eta \rightarrow 0} \frac{\gamma}{\eta} \int_{\mathcal{E}} F^T p \mu_{\gamma,\eta}(dq dp) = \lim_{\eta \rightarrow 0} \frac{1}{\eta} \int_{\mathcal{E}} [-F^T \nabla V(q) - \mathcal{L}_\gamma(F^T p)] \mu_{\gamma,\eta}(dq dp) \\
&= \beta \int_{\mathcal{E}} F^T p \mathcal{L}_\gamma^{-1} [F^T \nabla V(q) + \mathcal{L}_\gamma(F^T p)] \mu(dq dp) \\
&= |F|^2 + \beta \int_{\mathcal{E}} (F^T p) [p^T \nabla_q \mathcal{L}_{\text{ovd}}^{-1}(F^T \nabla V)] \mu(dq dp) + \frac{1}{\gamma} r_\gamma \\
&= |F|^2 + \int_{\mathcal{M}} (F^T \nabla_q^* \mathbf{1}) \mathcal{L}_{\text{ovd}}^{-1}(F^T \nabla V) \bar{\mu}(dq) + \frac{1}{\gamma} r_\gamma \\
&= |F|^2 + \beta \int_{\mathcal{M}} (F^T \nabla V) \mathcal{L}_{\text{ovd}}^{-1}(F^T \nabla V) \bar{\mu}(dq) + \frac{1}{\gamma} r_\gamma \\
&= |F|^2 + \bar{v}_F + \frac{1}{\gamma} r_\gamma,
\end{aligned}$$

where r_γ is uniformly bounded for $\gamma \geq 1$. This gives the desired result.

REMARK 4.1 The article Hairer & Pavliotis (2008) in fact studies the limiting behavior of the autodiffusion coefficient, as computed from (3.6):

$$\beta \bar{D}_F = \int_{\mathcal{M}} |F + \nabla_q \mathcal{L}_{\text{ovd}}^{-1}(F \cdot \nabla V)|^2 d\bar{\mu}.$$

Using $\mathcal{L}_{\text{ovd}} = -\beta^{-1} \nabla_q^* \nabla_q$, a simple computation shows

$$\begin{aligned}
\beta \bar{D}_F &= |F|^2 + 2 \int_{\mathcal{M}} F^T \nabla_q \mathcal{L}_{\text{ovd}}^{-1}(F \cdot \nabla V) d\bar{\mu} + \int_{\mathcal{M}} |\nabla_q \mathcal{L}_{\text{ovd}}^{-1}(F \cdot \nabla V)|^2 d\bar{\mu} \\
&= |F|^2 + 2 \int_{\mathcal{M}} (F^T \nabla_q^* \mathbf{1}) \mathcal{L}_{\text{ovd}}^{-1}(F \cdot \nabla V) d\bar{\mu} + \int_{\mathcal{M}} \nabla_q^* \nabla_q \mathcal{L}_{\text{ovd}}^{-1}(F \cdot \nabla V) \mathcal{L}_{\text{ovd}}^{-1}(F \cdot \nabla V) d\bar{\mu} \\
&= |F|^2 + \beta \int_{\mathcal{M}} (F^T \nabla V) \mathcal{L}_{\text{ovd}}^{-1}(F \cdot \nabla V) d\bar{\mu},
\end{aligned}$$

so that $\beta \bar{D}_F = |F|^2 + \bar{v}_F$.

4.12 Proof of Theorem 3.2

CASE $\alpha = 1$. Let us first consider the first order scheme $P_{\Delta t}^{\mathcal{C}, B + \eta \widetilde{\mathcal{L}}, A}$. Using the notation introduced in Section 4.3.2, and recalling the definition $B_\eta = B + \eta \widetilde{\mathcal{L}}$, we write

$$P_{\Delta t}^{\mathcal{C}, B + \eta \widetilde{\mathcal{L}}, A} = \text{Id} + \Delta t (A + B_\eta + \gamma \mathcal{C}) + \frac{\Delta t^2}{2} \mathcal{T} \left[(A + B_\eta + \gamma \mathcal{C})^2 \right] + \frac{\Delta t^3}{2} R_{\eta, \Delta t},$$

with

$$R_{\eta, \Delta t} = \int_0^1 (1 - \theta)^2 \mathcal{T} \left[(A + B_\eta + \gamma \mathcal{C}) P_{\theta \Delta t}^{\mathcal{C}, B + \eta \widetilde{\mathcal{L}}, A} \right]^3 d\theta.$$

Since

$$e^{\theta \Delta t B_\eta} - e^{\theta \Delta t B} = \eta \int_0^1 e^{\theta s B_\eta} \widetilde{\mathcal{L}} e^{\theta(1-s)B} ds,$$

it is easy to see that the operator $R_{\eta, \Delta t}$ can be rewritten as the sum of two contributions: $R_{\eta, \Delta t} = R_{0, \Delta t} + \eta \widetilde{R}_{\eta, \Delta t}$, where, for $\psi \in \mathcal{S}$, the smooth function $\widetilde{R}_{\eta, \Delta t} \psi$ can be uniformly controlled in η for $|\eta| \leq 1$. Finally, the evolution operator can be rewritten as

$$P_{\Delta t}^{\mathcal{C}, B + \eta \widetilde{\mathcal{L}}, A} = \text{Id} + \Delta t (\mathcal{L}_\gamma + \eta \widetilde{\mathcal{L}}) + \frac{\Delta t^2}{2} (\mathcal{L}_\gamma^2 + S_1 + \eta D_1) + \Delta t^2 \mathcal{R}_{\eta, \Delta t}, \quad (4.48)$$

where S_1 is defined in (4.18) (which corresponds to the case $\eta = 0$), $D_1 = (2\gamma C + B)\widetilde{\mathcal{L}} + \widetilde{\mathcal{L}}(2A + B)$, and

$$\mathcal{R}_{\eta,\Delta t} = \frac{\Delta t}{2} R_{0,\Delta t} + \frac{\eta \Delta t}{2} \widetilde{R}_{\eta,\Delta t} + \frac{\eta^2}{2} \widetilde{\mathcal{L}}^2.$$

We then compute

$$\begin{aligned} & \int_{\mathcal{E}} \left[\left(\frac{\text{Id} - P_{\Delta t}^{\gamma C, B + \eta \widetilde{\mathcal{L}}, A}}{\Delta t} \right) \varphi \right] (1 + \Delta t f_{1,0,\gamma} + \eta f_{0,1,\gamma} + \eta \Delta t f_{1,1,\gamma}) \, d\mu \\ &= - \int_{\mathcal{E}} \left[\left(\mathcal{L}_{\gamma} + \eta \widetilde{\mathcal{L}} + \frac{\Delta t}{2} (\mathcal{L}_{\gamma}^2 + S_1 + \eta D_1) + \Delta t \mathcal{R}_{\eta,\Delta t} \right) \varphi \right] (1 + \Delta t f_{1,0,\gamma} + \eta f_{0,1,\gamma} + \eta \Delta t f_{1,1,\gamma}) \, d\mu \\ &= - \eta \int_{\mathcal{E}} \left[\widetilde{\mathcal{L}} \varphi + (\mathcal{L}_{\gamma} \varphi) f_{0,1,\gamma} \right] \, d\mu - \Delta t \int_{\mathcal{E}} \left[\frac{1}{2} S_1 \varphi + (\mathcal{L}_{\gamma} \varphi) f_{1,0,\gamma} \right] \, d\mu \\ &\quad - \eta \Delta t \int_{\mathcal{E}} \left[\left(\widetilde{\mathcal{L}} \varphi \right) f_{1,0,\gamma} + \frac{1}{2} (\mathcal{L}_{\gamma}^2 + S_1) \varphi f_{0,1,\gamma} + (\mathcal{L}_{\gamma} \varphi) f_{1,1,\gamma} + \frac{1}{2} D_1 \varphi \right] \, d\mu \\ &\quad - \eta^2 \int_{\mathcal{E}} \left(\widetilde{\mathcal{L}} \varphi \right) (f_{0,1,\gamma} + \Delta t f_{1,1,\gamma}) \, d\mu - \frac{\Delta t^2}{2} \int_{\mathcal{E}} \left[(\mathcal{L}_{\gamma}^2 + S_1 + \eta D_1) \varphi \right] (f_{1,0,\gamma} + \eta f_{1,1,\gamma}) \, d\mu \\ &\quad - \Delta t \int_{\mathcal{E}} \mathcal{R}_{\eta,\Delta t} \varphi (1 + \Delta t f_{1,0,\gamma} + \eta f_{0,1,\gamma} + \eta \Delta t f_{1,1,\gamma}) \, d\mu. \end{aligned}$$

The first two terms in the last expression vanish by definition of $f_{0,1,\gamma}$ and $f_{1,0,\gamma}$, while the third one vanishes when the function $f_{1,1,\gamma}$ is defined by the Poisson equation

$$\mathcal{L}_{\gamma}^* f_{1,1,\gamma} = -\widetilde{\mathcal{L}}^* f_{1,0,\gamma} - \frac{1}{2} (\mathcal{L}_{\gamma}^2 + S_1)^* f_{0,1,\gamma} - \frac{1}{2} D_1^* \mathbf{1}. \quad (4.49)$$

It is easy to check that the right-hand side of this equation has a vanishing average with respect to μ (integrating with respect to μ and letting the adjoints of the operators acting on $\mathbf{1}$). The regularity of the functions is not an issue either since, by the results of Talay (2002) for instance (recalled in Section 4.3.1), the operator $\mathcal{L}_{\gamma}^{-1}$ is bounded on $\widetilde{H}^m(\mu)$ for any $m \geq 1$, so that the functions $f_{1,0,\gamma}$ and $f_{0,1,\gamma}$ in fact belong to all the spaces $\widetilde{H}^m(\mu)$. The right-hand side therefore is in \mathcal{H}^1 .

We then introduce the quasi inverse

$$\begin{aligned} Q_{\eta,\Delta t} &= -\mathcal{L}_{\gamma}^{-1} + \eta \mathcal{L}_{\gamma}^{-1} \widetilde{\mathcal{L}} \mathcal{L}_{\gamma}^{-1} + \frac{\Delta t}{2} [\text{Id} + \mathcal{L}_{\gamma}^{-1} (S_1 + \eta D_1) \mathcal{L}_{\gamma}^{-1}] \\ &\quad - \frac{\eta \Delta t}{2} \left(\mathcal{L}_{\gamma}^{-1} \widetilde{\mathcal{L}} \mathcal{L}_{\gamma}^{-1} (\mathcal{L}_{\gamma}^2 + S_1 + \eta D_1) \mathcal{L}_{\gamma}^{-1} + \mathcal{L}_{\gamma}^{-1} (\mathcal{L}_{\gamma}^2 + S_1 + \eta D_1) \mathcal{L}_{\gamma}^{-1} \widetilde{\mathcal{L}} \mathcal{L}_{\gamma}^{-1} \right), \end{aligned}$$

obtained by truncating the formal series expansion of the inverse operator by discarding terms associated with η^2 or Δt^2 . The quasi inverse is such that

$$\left(\frac{\text{Id} - P_{\Delta t}^{\gamma C, B + \eta \widetilde{\mathcal{L}}, A}}{\Delta t} \right) Q_{\eta,\Delta t} = \text{Id} + \eta^2 \mathcal{R}_{\eta,\Delta t}^1 + \Delta t^2 \mathcal{R}_{\eta,\Delta t}^2,$$

with $\mathcal{R}_{\eta,\Delta t}^2 = \mathcal{R}_{0,\Delta t}^2 + \eta \widetilde{\mathcal{R}}_{\eta,\Delta t}^2$. We then replace φ by $Q_{\eta,\Delta t} \psi$ and conclude as in Section 4.4.

CASE $\alpha = 2$. The result for the second order splitting is obtained by appropriate modifications of the proof written above for $p = 1$, similar to the ones introduced in Section 4.6. We will therefore mention only the most important point, which is the following. Replacing B by B_{η} in the expansion (4.25), we see

that

$$\begin{aligned} \frac{\text{Id} - P_{\Delta t}^{\gamma C, B_\eta, A, B_\eta, \gamma C}}{\Delta t} &= -\mathcal{L}_\gamma - \eta \widetilde{\mathcal{L}} - \frac{\Delta t}{2} (\mathcal{L}_\gamma + \eta \widetilde{\mathcal{L}})^2 - \Delta t^2 \left(\frac{1}{6} (\mathcal{L}_\gamma + \eta \widetilde{\mathcal{L}})^3 + S_2 + \eta \widetilde{S}_{2,\eta} \right) - \Delta t^3 R_{\eta, \Delta t} \\ &= -\mathcal{L}_\gamma - \eta \widetilde{\mathcal{L}} - \frac{\Delta t}{2} \mathcal{L}_\gamma^2 - \frac{\eta \Delta t}{2} (\mathcal{L}_\gamma \widetilde{\mathcal{L}} + \widetilde{\mathcal{L}} \mathcal{L}_\gamma) - \frac{\eta^2 \Delta t}{2} \widetilde{\mathcal{L}}^2 - \Delta t^2 \left(\frac{1}{6} \mathcal{L}_\gamma^3 + S_2 \right) \\ &\quad - \eta \Delta t^2 \left(\frac{1}{6} (\mathcal{L}_\gamma^2 \widetilde{\mathcal{L}} + \mathcal{L}_\gamma \widetilde{\mathcal{L}} \mathcal{L}_\gamma + \widetilde{\mathcal{L}} \mathcal{L}_\gamma^2) + \widetilde{S}_{2,0} \right) + \mathcal{R}_{\eta, \Delta t}, \end{aligned}$$

where $\mathcal{R}_{\eta, \Delta t}$ regroups operators of order $\Delta t^{3+\alpha} \eta^{\alpha'}$ or $\Delta t^{2+\alpha} \eta^{2+\alpha'}$ for $\alpha, \alpha' \geq 0$, the operator S_2 is defined in (4.24) and $\widetilde{S}_{2,\eta}$ satisfies

$$\begin{aligned} 12\widetilde{S}_{2,\eta} &= [A, [A, \widetilde{\mathcal{L}}]] - \frac{1}{2} [B, [\widetilde{\mathcal{L}}, A]] - \frac{1}{2} [\widetilde{\mathcal{L}}, [B, A]] + \gamma [\widetilde{\mathcal{L}}, [A+B, C]] + \gamma [A+B, [\widetilde{\mathcal{L}}, C]] \\ &\quad - \frac{\gamma^2}{2} [C, [C, \widetilde{\mathcal{L}}]] + \eta \left(\gamma [\widetilde{\mathcal{L}}, [\widetilde{\mathcal{L}}, C]] - \frac{1}{2} [\widetilde{\mathcal{L}}, [\widetilde{\mathcal{L}}, A]] \right). \end{aligned}$$

We next compute the dominant terms in

$$\int_{\mathcal{E}} \left[\left(\frac{\text{Id} - P_{\Delta t}^{\gamma C, B_\eta, A, B_\eta, \gamma C}}{\Delta t} \right) \varphi \right] (1 + \Delta t^2 f_{2,0,\gamma} + \eta f_{0,1,\gamma} + \eta \Delta t^2 f_{2,1,\gamma}) d\mu.$$

We consider only contributions of the form $\eta^\alpha \Delta t^{\alpha'}$ with $\alpha = 0, 1$ and $0 \leq \alpha' \leq 2$. The contributions in $\Delta t, \Delta t^2$ are the same as in the case $\eta = 0$ and therefore vanish. The contribution in η vanishes in view of the choice of $f_{0,1,\gamma}$. For the same reason, the contribution in $\eta \Delta t$ vanishes as well:

$$-\frac{\eta \Delta t}{2} \int_{\mathcal{E}} (\mathcal{L}_\gamma \widetilde{\mathcal{L}} + \widetilde{\mathcal{L}} \mathcal{L}_\gamma) \varphi + (\mathcal{L}_\gamma^2 \varphi) f_{0,1,\gamma} d\mu = -\frac{\eta \Delta t}{2} \int_{\mathcal{E}} (\mathcal{L}_\gamma \varphi) (\widetilde{\mathcal{L}}^* \mathbf{1} + \mathcal{L}_\gamma^* f_{0,1,\gamma}) d\mu = 0.$$

The contribution in $\eta \Delta t^2$ is proportional to

$$\int_{\mathcal{E}} \left[\left(\frac{\mathcal{L}_\gamma^2 \widetilde{\mathcal{L}} + \mathcal{L}_\gamma \widetilde{\mathcal{L}} \mathcal{L}_\gamma + \widetilde{\mathcal{L}} \mathcal{L}_\gamma^2}{6} + \widetilde{S}_{2,0} \right) \varphi + (\widetilde{\mathcal{L}} \varphi) f_{2,0,\gamma} + \left[\left(\frac{\mathcal{L}_\gamma^3}{6} + S_2 \right) \varphi \right] f_{0,1,\gamma} + (\mathcal{L}_\gamma \varphi) f_{2,1,\gamma} \right] d\mu.$$

The requirement that this expression vanishes for all functions $\varphi \in \mathcal{S}$ characterizes the function $f_{2,1,\gamma}$ (the discussion on the solvability of this equation following the same lines as the discussion on the solvability of (4.49)). The proof is then concluded as in the case $p = 1$.

4.13 Proof of Theorem 3.3

The proof of this result is obtained by modifying the proof of Theorem 2.9 presented in Section 4.9 by taking into account the nonequilibrium perturbation, as done in the proof of Theorem 3.2 presented in Section 4.12. We will therefore be very brief and only mention the most important modifications.

We write the proof for the scheme associated with the evolution operator $P_{\Delta t}^{\gamma C, A, B_\eta, A, \gamma C}$ for instance (since this is the case explicitly treated in Section 4.9 for $\eta = 0$). First, arguing as in Section 4.9, we see that it is possible to replace $P_{\Delta t}^{\gamma C, A, B_\eta, A, \gamma C}$ by

$$\pi P_{\text{ham}, \Delta t, \eta} \pi = \pi e^{\Delta t A/2} e^{\Delta t B_\eta} e^{\Delta t A/2} \pi$$

up to error terms in the invariant measure which are exponentially small in $\gamma \Delta t$. Note that $B_\eta = (F - \nabla V) \cdot \nabla_p$, so that the rules (4.34)-(4.35) are still valid. Therefore, introducing again $h = \Delta t^2/2$,

$$\begin{aligned} \pi P_{\text{ham}, \Delta t, \eta} \pi &= \pi + \frac{\Delta t^2}{2} \pi (A + B_\eta)^2 \pi + \frac{\Delta t^4}{24} \pi \left(A^4 + \frac{3}{2} A^2 B_\eta A + \frac{3}{2} A B_\eta A^2 + \frac{3}{2} B_\eta^2 A^2 + \frac{1}{2} B_\eta A^3 \right) \pi + \Delta t^6 R_{\Delta t, \eta} \\ &= \pi + h \pi \left(\mathcal{L}_{\text{ovd}} + \eta [\widetilde{\mathcal{L}}(A+B) + (A+B)\widetilde{\mathcal{L}}] + \eta^2 \widetilde{\mathcal{L}}^2 \right) \pi + \frac{h^2}{2} \left(\mathcal{L}_{\text{ovd}}^2 + D + \eta \widetilde{D}_1 + \eta^2 \widetilde{D}_2 \right) \pi \\ &\quad + \Delta t^6 R_{\Delta t, \eta}, \end{aligned}$$

where D is defined in (4.38), and the expressions of the operators \tilde{D}_i ($i = 1, 2$) are obtained by expanding the various terms $A^a B_\eta^b A^c$ in powers of η . Keeping only the dominant terms, we arrive at

$$\pi P_{\text{ham}, \Delta t, \eta} \pi = \pi + h \mathcal{L}_{\text{ovd}} \pi + \frac{h^2}{2} (\mathcal{L}_{\text{ovd}}^2 + D) + \eta h \pi \left[\tilde{\mathcal{L}}(A+B) + (A+B) \tilde{\mathcal{L}} \right] \pi + \frac{\eta h^2}{2} \tilde{D}_1 + \mathcal{R}_{\Delta t, \eta}.$$

Since

$$\pi \left(\tilde{\mathcal{L}}(A+B) + (A+B) \tilde{\mathcal{L}} \right) \pi = \pi \tilde{\mathcal{L}} A \pi = \tilde{\mathcal{L}}_{\text{ovd}},$$

we conclude

$$\pi P_{\text{ham}, \Delta t, \eta} \pi = \pi + h \left(\mathcal{L}_{\text{ovd}} + \eta \tilde{\mathcal{L}}_{\text{ovd}} \right) \pi + \frac{h^2}{2} \left(\mathcal{L}_{\text{ovd}}^2 + D + \eta \tilde{D}_1 \right) + \mathcal{R}_{\Delta t, \eta}.$$

This relation is the analogue of (4.48) in the overdamped limit, and the proof is carried on following the strategy presented in Section 4.9.

Acknowledgements

The authors thank Francis Nier for several fruitful discussions on the properties of Langevin-type generators. Ben Leimkuhler and Charles Matthews acknowledge the support of the Engineering and Physical Sciences Research Council (UK) and grant EP/G036136/1. G. Stoltz was partially supported by the project DYMOM (De la dynamique moléculaire, via l'homogénéisation, aux modèles macroscopiques de poroélasticité et électrocinétique) from the program NEEDS (Projet fédérateur Milieux Poreux MIPOR).

REFERENCES

- BOU-RABEE, N. & HAIRER, M. (2013) Nonasymptotic mixing of the MALA algorithm. *IMA J. Numer. Anal.*, **33**, 80–110.
- BOU-RABEE, N. & OWHADI, H. (2010) Long-run accuracy of variational integrators in the stochastic context. *SIAM J. Numer. Anal.*, **48**, 278–297.
- BOU-RABEE, N. & VANDEN-EIJNDEN, E. (2009) Pathwise accuracy and ergodicity of metropolized integrators for SDEs. *Commun. Pure Appl. Math.*, **63**, 655–696.
- BUSSI, G. & PARRINELLO, M. (2007) Accurate sampling using Langevin dynamics. *Phys. Rev. E*, **75**, 056707.
- CANCÈS, E., LEGOLL, F. & STOLTZ, G. (2007) Theoretical and numerical comparison of sampling methods for molecular dynamics. *Math. Model. Numer. Anal.*, **41**, 351–390.
- DEBUSSCHE, A. & FAOU, E. (2012) Weak backward error analysis for SDEs. *SIAM J. Numer. Anal.*, **50**, 1735–1752.
- ECKMANN, J.-P. & HAIRER, M. (2003) Spectral properties of hypoelliptic operators. *Commun. Math. Phys.*, **235**, 233–253.
- EVANS, D. J. & MORRIS, G. P. (2008) *Statistical mechanics of nonequilibrium liquids*. Cambridge University Press.
- GOODMAN, J. B. & LIN, K. K. (2009) Coupling control variates for Markov chain Monte Carlo. *J. Comput. Phys.*, **228**, 7127–7136.
- HAIRER, E., LUBICH, C. & WANNER, G. (2006) *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*. Springer Series in Computational Mathematics, vol. 31. Springer-Verlag.
- HAIRER, M. & MATTINGLY, J. C. (2011) Yet another look at Harris' ergodic theorem for Markov chains. *Seminar on Stochastic Analysis, Random Fields and Applications VI*. Progr. Probab., vol. 63. Birkhäuser/Springer, pp. 109–117.
- HAIRER, M. & PAVLIOTIS, G. (2008) From ballistic to diffusive behavior in periodic potentials. *J. Stat. Phys.*, **131**, 175–202.
- HASTINGS, W. K. (1970) Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, **57**, 97–109.
- HÉRAU, F. & NIER, F. (2004) Isotropic hypoellipticity and trend to equilibrium for the Fokker-Planck equation with a high-degree potential. *Arch. Ration. Mech. Anal.*, **171**, 151–218.
- JOUBAUD, R. & STOLTZ, G. (2012a) Nonequilibrium shear viscosity computations with Langevin dynamics. *Multi-scale Model. Sim.*, **10**, 191–216.
- JOUBAUD, R. & STOLTZ, G. (2012b) Nonequilibrium shear viscosity computations with Langevin dynamics. *arXiv preprint*, **1106.0633**.

- JOUBAUD, R. & STOLTZ, G. (2013) Erratum: Nonequilibrium shear viscosity computations with Langevin dynamics. *Multiscale Model. Sim.*, **11**, 411–414.
- KLOKOV, S. A. & VERETENNIKOV, A. Y. (2006) On mixing and convergence rates for a family of Markov processes approximating SDEs. *Random Oper. Stochastic Equations*, **14**, 103–126.
- KLOKOV, S. A. & VERETENNIKOV, A. Y. (2013) On local mixing conditions for SDE approximations. *Theor. Probab. Appl.*, **57**, 110–131.
- LEIMKUHLER, B. & MATTHEWS, C. (2013a) Rational construction of stochastic numerical methods for molecular sampling. *Appl. Math. Res. Express*, 34–56.
- LEIMKUHLER, B. & MATTHEWS, C. (2013b) Robust and efficient configurational molecular sampling via Langevin dynamics. *J. Chem. Phys.*, **138**, 174102.
- LELIÈVRE, T., ROUSSET, M. & STOLTZ, G. (2010) *Free-Energy Computations: A Mathematical Perspective*. Imperial College Press.
- MATTINGLY, J. C., STUART, A. M. & HIGHAM, D. J. (2002) Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stoch. Proc. Appl.*, **101**, 185–232.
- MATTINGLY, J. C., STUART, A. M. & TRETYAKOV, M. V. (2010) Convergence of numerical time-averaging and stationary measures via Poisson equations. *SIAM J. Numer. Anal.*, **48**, 552–577.
- MELCHIONNA, S. (2007) Design of quasisymplectic propagators for Langevin dynamics. *J. Chem. Phys.*, **127**, 044108.
- METROPOLIS, N., ROSENBLUTH, A. W., ROSENBLUTH, M. N., TELLER, A. H. & TELLER, E. (1953) Equations of state calculations by fast computing machines. *J. Chem. Phys.*, **21**, 1087–1091.
- MEYN, S. P. & TWEEDIE, R. L. (2009) *Markov Chains and Stochastic Stability (2nd edition)*. Cambridge University Press.
- PAZY, A. (1983) *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Applied Mathematical Sciences, vol. 44. New York: Springer.
- REY-BELLET, L. (2006) Ergodic properties of markov processes. *Open Quantum Systems II* (S. Attal, A. Joye & C.-A. Pillet eds). Lecture Notes in Mathematics, vol. 1881. Springer, pp. 1–39.
- SKEEL, R. D. & IZAGUIRRE, J. A. (2002) An impulse integrator for Langevin dynamics. *Mol. Phys.*, **100**, 3885–3891.
- STOLTZ, G. (2012) *Molecular Simulation: Nonequilibrium and Dynamical Problems*. Habilitation thesis. Université Paris Est. <http://tel.archives-ouvertes.fr/tel-00709965>.
- TALAY, D. (2002) Stochastic Hamiltonian dissipative systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Proc. Rel. Fields*, **8**, 163–198.
- TALAY, D. & TUBARO, L. (1990) Expansion of the global error for numerical schemes solving stochastic differential equations. *Stoch. Proc. Appl.*, **8**, 94–120.
- THALMANN, F. & FARAGO, J. (2007) Trotter derivation of algorithms for Brownian and dissipative particle dynamics. *J. Chem. Phys.*, **127**, 124109.
- TUCKERMAN, M. (2010) *Statistical Mechanics: Theory and Molecular Simulation*. Oxford University Press.
- VILLANI, C. (2009) Hypocoercivity. *Mem. Amer. Math. Soc.*, **202**.